

Aliasing of the Indian Ocean externally-forced warming spatial pattern by internal climate variability

S Gopika¹, Takeshi Izumo^{2,3}, Jérôme Vialard², Matthieu Engaigue^{2,3}, Iyyappan Suresh¹, M R Ramesh Kumar¹

¹ CSIR-National Institute of Oceanography, Goa, India.

² LOCEAN, IPSL, Sorbonne Universités (UPMC, Univ. Paris 06) -CNRS-IRD-MNHN, Paris, France

³ Indo-French Cell for Water Sciences, IISc-NIO-IITM-IRD Joint International Laboratory, NIO, Goa, India
Submitted to Climate Dynamics on 01 July 2019

Corresponding author:

Iyyappan Suresh(isuresh@nio.org)
Physical Oceanography Division
CSIR - National Institute of Oceanography
Dona Paula – 403004, Goa, India

Abstract

Coupled Model Inter comparison Project (CMIP5) models project an inhomogeneous anthropogenic surface warming of the Indian Ocean by the end of the 21st century, with strongest warming in the Arabian Sea and Western equatorial Indian Ocean. Previous studies have warned that this “Indian Ocean Dipole (IOD)-like” warming pattern could yield more Arabian Sea cyclones, more extreme IOD events and decrease monsoonal rains. Here we show that CMIP5 models also produce an “IOD-like” pattern over the 1871-2016 period, in broad agreement with observations. Single-models ensemble simulations however indicate a strong aliasing of the warming pattern “signal” by the internal climate variability “noise” over that period. While the average Indian Ocean warming emerges around 1950 in CMIP5 and observations, regional contrasts are more difficult to detect. The only detectable signal by 2016 in CMIP5 is a stronger Arabian Sea than Bay of Bengal warming in > 80% of the models, which is not detected in HadSST3 observations. Conversely, observations already detect a stronger Northern than Southern Indian ocean warming, while this signal only emerges by ~2060 in > 80% of the models. Subsampling observations to only retain the most accurate values however indicate that this observed signal most likely results from sampling issues in the Southern hemisphere. In light of this large aliasing by internal climate variability and observational uncertainties, the broad agreement between CMIP5 and observations over 1871-2016 may be largely coincidental. Overall, these results call for extreme caution when interpreting spatial patterns of anthropogenic surface warming.

Key words: Indian Ocean, Anthropogenic climate change, Natural climate variability, Sea Surface Temperature (SST), Coupled Model Inter comparison Project (CMIP), spatial pattern of Indian Ocean SST change.

1. Introduction

The Indian Ocean (IO), which gathers one third of the Earth's population around its rim, has warmed steadily over the past century, with a faster warming rate since the 1960s (Fig. 1a; Alory et al. 2007; Alory and Meyers 2009; Du and Xie 2008; Levitus et al. 2009; Han et al. 2014). This IO-averaged warming rate is broadly consistent across historical simulations from the Coupled Model Intercomparison Project (CMIP; Taylor et al. 2012) and observational products (Fig. 1a; Alory et al. 2007; Du and Xie 2008; Alory and Meyers 2009; Dong et al. 2014; Yao et al. 2016). It is largely attributable to the anthropogenic forcing rather than the natural external forcing (volcanic and solar) variations that have much weaker effect (e.g. Dong et al. 2014; Dong and Zhou 2014). The IO warming is projected to further increase in the course of the 21st century in response to unabated greenhouse gases emissions (Fig. 1a). This IO-averaged warming is thought to induce worldwide impacts, including an intensification of the Walker cell (e.g. Luo et al. 2012; Hamlington et al. 2014), more frequent droughts in the Mediterranean region (e.g. Hoerling et al. 2012) and food security issues in the eastern and southern Africa (e.g. Funk et al. 2008).

This robust IO warming is however far from being spatially homogeneous in both models (Ihara et al. 2009; Zheng et al. 2013; Liu et al. 2015; Yao et al. 2016) and observations (Alory et al. 2007; Ihara et al. 2008; Han et al. 2014; Swapna et al. 2014; Roxy et al. 2015; Rahul and Gnanaseelan 2016). CMIP5 projections indeed indicate a larger warming in the western than in the eastern tropical Indian Ocean by the end of the 21st century (e.g. Zheng et al. 2013; Dong et al. 2014), a pattern of Sea Surface Temperature (SST) change that is often referred to as “IOD-like” in reference to the typical SST anomalies observed during positive Indian Ocean Dipole events (IOD; e.g. Reverdin et al. 1986; Saji et al. 1999; Webster et al. 1999; Murtugudde et al. 2000). The CMIP5 Multi-Model Mean (hereafter, MMM) also indicates a stronger warming in the western equatorial IO and Arabian Sea over the 1871-2016 historical period (Fig. 1b), *i.e.* a warming relative to the IO mean in the northwest and cooling relative to the IO mean in the Bay of Bengal and southern subtropics. Published literature did not reach a consensus on which part of the IO has been warming the most in observations. This may partly be due to the fact that the pattern of SST change (hereafter, pSSTc) is very sensitive to the period over which it is computed, as demonstrated by the clear differences between SST changes computed over the last 50 years (Fig. 1c) and over the 1871-2016 period (Fig. 1d).

Several studies however suggested that the exact pSSTc has important consequences for regional climate impacts. The pattern of projected SST changes relative to the tropical mean (also known as the “relative SST”) for instance matches well with that of tropical cyclones maximum

potential intensity changes (Vecchi and Soden 2007), providing a plausible explanation for the observed cyclogenesis increase in the Arabian Sea and decrease in the Bay of Bengal (Murakami et al. 2013; Murakami et al. 2017). Enhanced warming over the western equatorial IO (Kulkarni, 2012; Roxy et al. 2014; Roxy et al. 2015) and inter hemispheric SST gradient changes (Saha et al. 2014) are both proposed as plausible explanations for the observed decreasing tendency in Indian monsoon rainfall over the last three decades. Changes in the zonal SST gradient along the equator are also thought to be conducive to a future increase of the frequency of extreme positive IOD events, a major source of regional climate hazards (Cai et al. 2014). It is hence important to know if we can trust projections of the pSSTc, as they have important implications for the future changes in the summer monsoon, tropical cyclones, and occurrence of extreme IOD events.

When computed over the entire historical record, the observed IO trend pattern displays an enhanced warming over the Arabian Sea and equatorial IO and a reduced warming over the southwestern IO (Fig. 1d; e.g. Ihara et al. 2008; Roxy et al. 2015). This observed 1871-2016 pSSTc shares similarities with that of the CMIP5MMM (Fig. 1b), including an enhanced warming in the western equatorial IO and Arabian Sea (Fig. 1b,d). It however also displays noticeable differences, such as a stronger warming in the eastern equatorial IO and a weaker warming in the southern subtropics than in the CMIP5 MMM (Fig. 1b,d). These differences can arise from three factors: (1) aliasing by internal climate variability (averaged out by taking the mean of 36 CMIP5 models), (2) observational uncertainties and/or (3) model biases. There is indeed a strong aliasing of the externally-forced SST trends by the natural climate variability in tropical regions (Deser et al. 2012), particularly by natural decadal to multi-decadal variability that has similar timescales to changes induced by anthropogenic forcing. For instance, in the tropical Pacific, the Interdecadal Pacific Oscillation transition to a negative phase during the last decades has strongly contributed to reduce the long-term warming tendency in the central and eastern equatorial Pacific (e.g. Tokinaga and Xie 2012). In the IO, the completely different pSSTc obtained when considering two different periods (Fig. 1cd) could hence be attributable to such internal variability. Alternatively, the very sparse observational coverage before the satellite era in the IO (e.g. Terray 1994; Deser et al. 2010; Izumo et al. 2014) could potentially yield large uncertainties in the pre-1980 SST estimates and hence in the pSSTc. Finally, CMIP5 projections themselves might be erroneous. Li et al. (2016) for instance argue that the “IOD-like” pSSTc is largely spurious, and can be linked to an overestimated IOD variability in CMIP5 models.

Given the strong potential implications of the pSSTc for regional climate impacts, it is important to evaluate whether observations and models already allow us to make a clear statement on

which parts of the IO warm most. In the present paper, we aim at assessing whether the externally-forced pSSTc is already detectable in the IO and whether its aliasing by internal climate variability can explain some of the differences between models and observations. We will use climate model simulations and the concept of “emergence time” (defined as the time when the climate change “signal” irreversibly emerges from the background climate “noise”, Hawkins and Sutton 2012) to assess if/when SST gradients changes induced by external forcing become detectable. We will also describe observational issues in the IO and investigate whether subsampling the models and observations to only retain grid points with the most accurate observations can reduce apparent discrepancies.

The paper is organized as follows. Section 2 describes the observations, CMIP5 simulations and the methodology that we use in this paper. In section 3, we will first demonstrate the strong aliasing of the externally-forced pSSTc by the internal climate variability in historical 10-members ensemble simulations. We will then apply the emergence time concept to the CMIP5 database (using the historical and RCP8.5 unmitigated scenario). We will specifically consider the IO-averaged SST, but also SST gradients between the Arabian Sea and Bay of Bengal, between the west and east equatorial IO, and between the north and south tropical IO, in relation to their relevance for changes in tropical cyclones, IOD extreme events and monsoons. In section 4, we will discuss observational issues and assess how they affect data/model comparison. Finally, we will summarize our results in section 5 and discuss whether the “IOD-like” CMIP5 MMM warming pattern of Figure 1b, d can be considered as realistic or not.

2. Data and methods

We will mostly use the HADSST3 (Hadley Centre SST) observational product (Kennedy et al. 2011a, 2011b). This data set is built from the International Comprehensive Ocean-Atmosphere Dataset (ICOADS) that combines measured SSTs from ships, oceanographic stations, moored buoys, drifting buoys and research vessels (Woodruff et al. 2010). These data are averaged into a $5^\circ \times 5^\circ$, monthly bins, after applying quality-check and bias adjustments to reduce spurious trends caused by changes in measuring techniques. We chose this dataset for two main reasons. First, it provides the number of observations and an estimate of the measurement and sampling uncertainty in each grid cell, which we need for evaluating the influence of observational sampling. Second, compared to other well-known observational datasets such as ERSST (Huang et al. 2015) or HadISST (Rayner et al. 2003), HADSST3 does not use any interpolation method to fill gaps in data-sparse regions, hence avoiding the spurious generation of trends in region for which we have little data. While having gaps, this dataset is hence the closest to the original *in-situ* data. To assess the robustness of the results

obtained with this dataset, we will compare results obtained with HADSST3 with those obtained with the other observational products in the discussion section. These products are ERSST-V3b (Smith et al. 2008), ERSST-V4 (Huang et al. 2015), ERSST-V5 (Huang et al. 2017), COBE-SST (Folland and Parker 1995), COBE-SST2 (Hirahara et al. 2014) and KAPLAN-SST (Kaplan et al. 1998). All of these products are available over the 1871-2016 period, except the two COBE products, which are only available over 1891-2016. We will thus compute the trends over the 1871-2016 period for all products except COBE. Additional analyses (not shown) however indicate that the observational trends are very similar when computed either over 1871-2016 or 1891-2016, indicating that any difference between the COBE and other products arises from the data processing, not from the trend computation period.

We will also analyze simulations from 36 models in the CMIP5 database (see Table 1; same models as Parvathi et al. 2017). We combined simulations forced by the observed external (i.e. anthropogenic and natural) forcing over the historical period (1871-2005) with projections for 2006-2100 under the RCP8.5 scenario (Riahi et al. 2011). RCP8.5 is an unmitigated scenario with radiative forcing reaching 8.5 W.m^{-2} relative to the pre-industrial conditions by 2100. We chose this scenario for two reasons. First, it is the strongest forcing available and will give a lower-bound estimate of the emergence time of the IO pSSTc. Second, in a recent agreement to reduce emissions, countries have set their emission limit close to the current levels (and close to the RCP8.5 scenario) until at least 2030 (Rogelj et al. 2016). When several members (i.e. starting from different initial conditions) are available for a given model, we analyzed the first member, in order to give equal weight to each model.

Models with available 10-members ensembles allow us evaluating the influence of internal climate variability. Indeed, due to the chaotic nature of climate, different members will sample completely different realizations of the internal climate variability, including phenomena like the El Niño-Southern Oscillation or the Interdecadal Pacific Oscillation, a few years after the start of the simulation. This will hence allow us to evaluate how the internal climate variability “noise” (i.e. the spread amongst various members) compares with the externally-forced “signal” (estimated from the ensemble mean). Our CMIP5 database includes two models for which 10-member ensembles are available: CNRM-CM5 and CSIRO-Mk3-6-0. Although not included in our 36 CMIP5 models database due to the unavailability of the corresponding RCP8.5 runs, two other models (HADCM3 and GFDL-CM2p1) also have 10-member ensembles available for the historical period. All the available monthly SST datasets (both simulations and observations) were re-gridded to a common $2^\circ \times 2^\circ$ grid.

We selected four indices to assess the externally-forced SST changes in the IO. The first one is simply the IO-averaged (30°E to 120°E and 25°S to 25°N) SST anomalies (relative to the 1871-2016 mean seasonal cycle). As explained in the introduction, the Arabian Sea / Bay of Bengal, west / east equatorial IO and interhemispheric SST gradient changes all have potentially important climatic consequences. We hence defined three indices to quantify these SST gradient changes. The AB (for Arabian Sea minus Bay of Bengal) index is defined as the averaged SST anomalies within 40°E to 80°E, 5°N to 25°N minus those within 80°E to 100°E, 5°N to 25°N, the WE (West minus East equatorial IO) index within 40°E to 80°E, 5°S to 5°N minus those within 80°E to 105°E, 5°S to 5°N, and the NS (North minus South IO; interhemispheric) index within 40°E to 100°E, 5°N to 25°N minus those within 50°E to 110°E, 25°S to 10°S (see Fig. 1b,c,d for an illustration of the various averaging boxes listed above).

We estimate the SST change over a given period as follows. For a given time series of SST T_i at time t_i , we estimated the trend a and the interception b through a least-square linear regression as follows (ϵ_i being the residuals of the regression):

$$T_i = a t_i + b + \epsilon_i \quad (1)$$

The SST change was then estimated over the entire length of the time series $\Delta T = a \Delta t$ (in °C). Note that this estimate of the SST change is independent of the reference period for computing the SST anomalies (e.g. 1871-2016 on Fig. 1a), and refers to the change between the start and the end of the time series. We preferred this method over that of simply taking the SST difference between the beginning and the end of the time series (e.g. Parvathi et al. 2017), as the least-square regression allows a more efficient filtering of the interannual/decadal variability in the time series. Obtaining a trend through a regression is sometimes not optimal for noisy datasets, and in such cases a robust Theil-Sen trend estimator may prove to be useful (e.g. Yue and Pilon 2004). We however verified that the trend maps computed with the linear regression and Theil-Sen methods are very similar in both observations and models (not shown), and we hence preferred to use the more computationally-efficient least-square method throughout the paper. The 90% confidence intervals on the SST changes are obtained using a two-tailed t-test, with the effective number of degrees of freedom estimated based on formula number 30 in Bretherton et al. (1999). Using the bootstrapping procedure of Yue and Pilon (2004) to estimate the 90% confidence interval yields nearly identical results (not shown).

Hawkins and Sutton (2012) introduced the concept of emergence time as the time when the externally-forced “signal” irreversibly emerges from the climate “noise” (Hegerl et al. 2006). When considering CMIP5 simulations or observations (for which we have a single realization), we use the

90% confidence interval of the trend as a measure of the climate noise. We then define the emergence time as the last time in the time series when the trend of a given index becomes significantly positive at the 90% confidence level. This is illustrated on Figure 2. The upper panels display the AB index time series for a given model (CSIRO-MK3-6-0) and observational dataset (HadSST3) while the lower panels show the corresponding SST change, estimated as explained above, over a period starting in 1871 with an end point varying between 1920 and 2100 (2016 for observations). The confidence interval (Fig. 2c,d) largely reflects uncertainties in the long-term trend associated with climate-related fluctuations in the AB index (Fig. 2a,b). The AB index remains significantly positive at the 90% confidence level in the CSIRO-MK3-6-0 from 2008 onward: 2008 thus is the AB index emergence time for that model. In HadSST3, the AB index has a negative trend until an end point ~ 1940 , with a wide 90% confidence interval. The trend then becomes marginally positive in the 90% confidence limit around 1965, but then decreases and is not significantly positive till date: this indicates that a positive AB index has therefore not yet emerged in HadSST3.

3. Projected SST gradient changes in CMIP5

In this section, we will first demonstrate the strong aliasing of the IO pSSTc by internal climate variability, using the four 10-member historical ensembles introduced in section 2. We will then turn to the entire CMIP5 database to estimate emergence times for the four SST indices discussed above (IO, AB, WE and NS) in section 3.2.

3.1 Aliasing by internal climate variability

As discussed in section 2, the internal climate variability should be largely uncorrelated between various members of the same model. The ensemble mean of a given model hence gives an estimate of the externally-forced signal in that model (averaging 10 uncorrelated climate “noise” realizations should reduce the amplitude of the noise by a factor of $10^{1/2} \approx 3$). The pSSTc spread amongst various members of a single model ensemble can then be entirely attributed to the model’s internal variability (unlike in the CMIP5 ensemble, where the different model physics also contribute to the spread between models).

Figure 3 illustrates that the internal variability can yield very different long-term SST changes in various members of the same model, even over the relatively long 1871-2005 period. Let us first focus on the inter-member differences in each model rather than on inter-model differences (which will be discussed comprehensively by considering the entire CMIP5 database in section 3.2). The left column displays the 1871-2005 SST change ensemble mean map, i.e. an estimate of the pSSTc for

each model. The middle and right columns show the SST change for the two ensemble members with the lowest pattern correlation (i.e. the most dissimilar patterns of SST change). For all models, the SST change of an individual member can either look somewhat alike the ensemble mean (middle column, with pattern correlations to the ensemble mean between 0.32 and 0.81) or quite different (right column, with pattern correlations to the ensemble mean between 0.02 and 0.24). The most extreme case of aliasing occurs for the CSIRO model, for which two members display almost opposite SST change maps (Fig. 3e,f; pattern correlation of -0.76). This may be related to a weak externally-forced pSSTc in that model, as illustrated by the small ensemble-mean response relative to other models on Figure 3d. But any of the models can yield quite different maps of the 1871-2005 SST trends (with minimum pattern correlation between the members of -0.13, -0.13 and -0.3 for the three other models). This suggests that internal climate variability can strongly alias the externally-forced pSSTc, even over the extended 1871-2005 observational period. This has strong consequences for comparing SST changes between models and observations. Besides the potential model error issues, comparing one model realization with observations is indeed analogous to comparing two individual members: different realizations of the internal climate variability in each can give rise to very different spatial pSSTc.

The 90% significance of the trend is also indicated with black dots for individual members on the middle and right columns of Figure 3. Interestingly, it shows that the trend of an individual member can be statistically significant in a region where there is little agreement between individual members, *e.g.* for HadCM3 east of Madagascar (Fig. 3g,i) or GFDL in the south-eastern tropical IO (Fig. 3j,k), suggesting that internal variability can contribute to the 1871-2005 SST change (i.e. includes some centennial signals). This is a testimony to the strong aliasing of the pSSTc by the internal climate variability, even over the entire period of the observational record.

Figure 4 displays the distribution of the 1871-2005 changes of the IO-averaged SST and three other gradient indices introduced in section 2 (AB, WE, NS) from the four 10-member ensembles and from all CMIP5 models listed in Table 1. For the IO-averaged warming, the four ensembles display quite different climate sensitivities, but all indicate a median 1871-2005 SST warming between +0.35 and +0.8°C, with a relatively weak ensemble spread (typically less than $\pm 0.1^\circ\text{C}$). I.e. the aliasing of the IO-average SST warming rate by internal climate variability is quite weak over the entire observational period, and inter-model differences are dominated by different climate sensitivities. In contrast, all gradient indices display a much larger impact of the internal variability on the index change. The ensemble spread (“noise”) tends to be larger (up to $\pm 0.2^\circ\text{C}$), while the signal (estimated from the median) is always far smaller than for the IO-averaged warming. As a result, a

lot of the ensemble distributions intersect the zero axis (two models for the AB, three for the NS and four out of the four models for the WE gradient). This indicates that the aliasing by internal climate variability prevents the detection of the sign of most SST gradient changes over the historical period in these models. This larger aliasing of the SST gradient indices compared to the IO-averaged index could be related to two factors. First, the spatial contrasts in warming captured by the gradient indices are much smaller than the basin-averaged warming (Fig. 4), *i.e.* the signal is smaller for the gradient changes than for the IO-averaged index. Second, some of the decadal climate variability (e.g. decadal IOD modulation) will induced opposite SST signals in various regions of the basin. While the IO-averaging will tend to reduce this signature, the difference between various regions can emphasize it, *i.e.* the noise can be larger for gradients than for the IO-averaged index.

Figure 4 also includes the distribution of the trends for the four indices, obtained from the 36 CMIP5 single-member simulations. This distribution suggests detectable or almost detectable changes for the AB index (with ~90% of the models indicating a positive AB change by 2005), but not for the two other gradient indices (for which ~60% of the models agree at best). We will evaluate this more precisely using the emergence time concept in section 3.2, but it is interesting to point that the CNRM and CSIRO 10-member ensembles sometimes yield, for some of their members, index change values that are outside of the full CMIP ensemble. This indicates that, over 1871-2005, the CMIP5 spread in pSSTc may be underestimated, because we only consider a single member per model and hence insufficiently sample aliasing by internal climate variability. This will be addressed by the emergence time approach of section 3.2, which considers the influence of internal variability.

3.2 Emergence time of climate change SST patterns from CMIP5

In the previous section, we mainly considered historical ensemble experiments for four different models to illustrate the strong aliasing of pSSTc by internal climate variability over the historical period. In this section, we aim at assessing whether the enhanced warming is already detectable in any region of the IO from the entire CMIP5 historical and RCP8.5 database. While each of the ensembles in section 3.1 only sampled the influence of internal climate variability, the 36-models CMIP5 database will also sample effects associated with differences in model physics, which can yield contrasted responses to external radiative forcing.

Figure 5 displays the MMM SST change relative to the tropical IO mean change over the entire period. While qualitatively similar to the 1871-2016 one (Fig. 1b), the 1871-2099 MMM pSSTc exhibits a much larger amplitude, as expected. This figure suggests a remarkable inter-model agreement, with more than 90% of the models indicating that the Arabian Sea and western equatorial

IO warm more, and that the Bay of Bengal and Southern IO warm less than the IO average (contours on Fig.5). The green curve on Figure 5a further quantifies this inter-model agreement by displaying the pattern correlation between the 1871-2099 pSSTc in the 36 individual CMIP5 models and the MMM displayed on Figure 5b (models are ordered based on decreasing pattern correlation). There is a very good agreement between the long-term SST change simulated by individual CMIP5 models over this long period, with pattern correlation to the MMM above 0.8 for about half of the models and larger than 0.5 for all models but one. This strong consistency between various models implies that the aliasing of pSSTc by internal climate variability is weak over the 1871-2099 period. This is due to a better filtering of internal climate variability (“noise”) over this long period (230 years), and to a much stronger signal due to the very strong post-2005 radiative forcing in RCP8.5. The main factors differentiating various models over this long period are hence most likely the different physical parameterizations and resolutions (e.g. Dommenges 2012).

The MMM pSST cover the historical period has a weaker amplitude but rather similar spatial pattern to the one computed over the 1871-2099 period (Fig. 1b vs. Fig. 5b), except for the southern IO cooling relative to the IO mean, which is nearly absent over the historical period. The excellent inter-model pSSTc agreement over the long 1871-2099 period (green curve on Fig. 5a) is however considerably weaker over the historical period (blue curve on Fig. 5a; about 1/3 of pattern correlations below 0.5). This is also the case when focusing only on models with a very consistent SST change over 1871-2099 (e.g. the 18 models having a pattern correlation with the MMM > 0.8), i.e. when focusing on models with a very consistent response to anthropogenic forcing. In line with section 3.1, this suggests that internal climate variability is a major source of uncertainty on the 1871-2005 pSSTc, as a shorter period leads to a weaker filtering of the internal climate “noise” and to a weaker radiative forcing. The pattern correlation distribution for the two CMIP5 models with 10-member ensembles can also vary a lot depending on the member (Fig. 5a, 0.12 to 0.65 for CSIRO and -0.12 to 0.57 for CNRM). This illustrates again that a non-negligible component of the inter-model spread in historical SST trend can be attributed to internal climate variability.

Figure 6 displays time series of annual-mean values of the three SST gradient indices for CMIP5 models. The MMM (thick black line) clearly reveals an upward trend for the three indices. On the other hand, individual model trajectories (light grey curve, with CNRM-CM5 and CSIRO-MK3-6-0 highlighted in red and blue, respectively) exhibit large interannual and decadal variations that make the detection of trends difficult. This is particularly the case for the CNRM-CM5 model (red curve), for which the trend (the signal) is weak, and difficult to detect behind the climatic “noise”, except maybe at the end of the record for the NS index. The upward trend is much more

visible for the three indices in the CSIRO-MK3-6-0 model (blue curve), but yet difficult to detect before around 2000. Figure 6 hence qualitatively supports the notion that a lot of the changes in surface temperature gradients are difficult to detect, due to aliasing by internal variability. Let us now turn to the more quantitative emergence time concept to evaluate when trends in the three indices become detectable in individual models.

Figure 7 displays the emergence time distribution (bar plot) and its cumulated values (dashed black curve) for the IO-averaged SST and the three gradient indices of the 36 CMIP5 models. Figure 7a indicates that the IO-averaged warming is already detectable by 1940 in ~90% and by 2000 in 100% of the models, *i.e.* the IO-averaged warming is already detectable in the CMIP5 database, as already reported by many studies (*cf.* introduction). Figure 7bcd indicates that most CMIP5 models agree on a warmer Arabian Sea and western equatorial Indian ocean, and cooler southern subtropics relative to the IO average by 2100, *i.e.* positive AB, WE, and NS indices by 2100 in most models (~90% to 100%; Fig. 7bcd). However, while the IO average warming emerges before 2000 in all models, this is not the case for any of the SST gradient indices. This underlines again that changes in SST gradients are more difficult to detect than the IO-averaged warming trend. In 2020 (*i.e.* ~ now, at the time of writing), a stronger warming signal in the Arabian Sea than in the Bay of Bengal has already emerged in 80% of the models, and this proportion rises to 100% by 2080. Only 40% of the models indicate a warmer western than eastern equatorial IO in 2020, and 90% by 2100. A stronger warming of the northern relative to the southern tropical IO is only detectable in 20% of the models in 2020, but this number quickly rises to 60% in 2040 and 90% in 2080. We will consider that a gradient change is detected in the CMIP5 database when it is detected for > 80% of the models. With this definition, the AB gradient change is already detectable, but the NS and WE gradients changes will not be detectable before ~2060 and ~2080, respectively.

We expect the emergence time to result from a competition between the externally-forced “signal” and the internal climate “noise”. To verify this, we estimate the “signal” for each model and index as the 1871-2099 trend (as described in section 2), and the “noise” as the standard deviation of departure from the linear trend estimate. Table 2 provides the partial correlations between the emergence time and the signal / noise, for the IO, AB, WE and NS indices. The partial correlation between the emergence time and the “signal” (excluding the effect of the “noise”) is always significantly negative at the 95% level (-0.37 to -0.77), confirming that the larger the signal to be detected, the earlier it will emerge. The partial correlation between the emergence time and “noise” (excluding the effect of the “signal”) is on the other hand always significantly positive at the 95% level (0.39 to 0.57): the larger the noise, the later the signal will emerge. All partial correlations of

table 2 being statistically significant, the signal and noise both matter for setting the emergence time of all indices.

4. Are changes in CMIP5 SST gradients consistent with observations?

We have so far focused on models to demonstrate the strong aliasing of the pSSTc by internal climate variability. But data coverage also strongly affects our ability to estimate and detect long-term SST changes in observations. In this section, we will first describe *in-situ* data coverage, and then compare CMIP5 models with *in-situ* observations by using a common subsampling for both.

Figure 8 assesses the evolution of the *in-situ* data coverage since 1871. The data coverage has considerably increase dsince 2000, coincident with the advent of Argo floats (Fig. 8d). Prior to 2000, the regional data density generally decreases in the following order: Arabian Sea, Bay of Bengal, eastern equatorial IO, western equatorial IO and southern tropical IO. The high data coverage in the Arabian Sea and Bay of Bengal is mainly related to two shipping routes crossing the southern end of these basins, which also account for a large fraction of IO data before the 1960s (Fig. 8b,c). The AB index is thus probably the gradient index that is best constrained by observations. The very poor sampling of the NS southern pole and poor sampling of the AB western pole on the other hand results in poor observational constraints of these two gradients, especially before the 1960s. This situation is made worse by the three large temporal gaps coinciding with the two world wars and the Kippur war / seventies oil shock. The strong aliasing by internal climate variability described earlier and the poor data coverage hence call for cautiousness when interpreting long-term trends estimated from such poorly-sampled observational record.

As mentioned in the introduction, the observed pSSTc derived from HadSST3 over the full historical period (Fig. 1d) shares some resemblance with the CMIP5 MMM simulate dpSSTc (Fig. 1b), with a larger warming in the western equatorial IO and Arabian Sea than in the Bay of Bengal and southern subtropics. The amplitude of the signals are however systematically two to three times larger in observations, which for instance exhibit a much stronger cooling (relative to the IO average) in the southern hemisphere than CMIP5 models. These apparent discrepancies between models and observations could partly be attributable to observational sampling issues. It is indeed puzzling to see that the observed pSSTc strongly echoes that of the spatial coverage (Fig. 8a-c vs. Fig. 1d), the maximum warming in the southern Arabian Sea and eastern equatorial IO being collocated with to the two shipping routes in the southern parts of the Arabian Sea and the Bay of Bengal while the maximum relative cooling in the southern IO occurs in a very poorly sampled region.

To investigate the impact of the data coverage, we have thus subsampled the models, by only retaining the grid cells where HadSST3 provides an SST estimate. Some HadSST3 grid cells are built from a very limited number of data, and hence subject to large representation errors. We have thus further subsampled HadSST3 and the models consistently by masking grid cells where the HadSST3 observational error estimate exceeds a threshold of either 0.5, 0.4 or 0.3°C. The number of valid grid cells that remain after such a subsampling is provided in Table 3. While 60-90% of grid points provide SST estimates in the original HADSST3 dataset, this percentage drops to 30-40% for the most severe subsampling in the two boxes involved in the AB index. The WE and NS indices are both computed using one box (the western equatorial and southern IO, respectively) for which the percentage of valid grid points drops to 5-10%. This illustrates again the poor observational constraints on the NS and WE gradients relative to AB.

Figure 9 compares the distribution of the changes in the 4 indices in CMIP5 with those in HadSST3, with different levels of subsampling. The distribution of the modelled IO, AB and WE indices are not much influenced by subsampling (Fig. 9a,b,c). In contrast, the modelled NS index distribution shifts upward as subsampling is increased (Fig 9d). This suggests that the poor observational coverage in the Southern IO emphasizes the apparent faster warming of the Northern hemisphere. Subsampling has a larger impact on the observed indices, suggesting a relatively large influence of representation errors on the observational estimate of long-term trends. This is in particular the case for the NS gradient (Fig. 9d), due to the very limited sampling in the southern hemisphere. Subsampling improves the agreement between the modelled distribution and HadSST3 for all indices except AB, for which HadSST3 always remains at the lower edge of the model distribution (Fig.9b).

Figure 9 allows discussing which index changes are detectable in observations over 1871-2016. Based on our definition, a signal is indeed detectable when it is different from zero in the 90% confidence limit (i.e. when the green shading does not intersect the zero axis on Fig. 9). The IO-average warming is detectable in observations (in fact before 1950, in close agreement with models; not shown), irrespective of the subsampling. Observed changes in the AB and WE indices are on the other hand not detectable, irrespective of the subsampling. Observations suggest a detectable stronger warming in the Northern hemisphere, except when the strongest subsampling is applied, and this signal is at the edge of detection (Fig. 7d).

To summarize, the 1871-2016 observed and modelled SST changes are consistent for the WE index, and both suggest undetectable changes by 2016. Models also suggest that changes in the NS gradient are currently not detectable, while observations suggest a stronger northern hemisphere

warming. This stronger warming is however reduced and at the marge of detection when considering observational uncertainties. Finally, models indicate a detectable stronger warming of the Arabian Sea than of the Bay of Bengal, while observations do not, irrespective of the subsampling.

5. Summary and discussion

5.1 Summary

The pattern of Sea Surface Temperature change (pSSTc) associated with anthropogenic climate change has important implications for regional climate impacts in the Indian Ocean (IO). Here we focus more specifically on the interhemispheric (NS), western minus eastern equatorial IO (WE) and Arabian Sea minus Bay of Bengal (AB) SST gradients, as they notably matter for future changes in the monsoon, extreme IOD events and tropical cyclones.

The CMIP5 Multi-Model Mean (MMM) projections for 2100 under the RCP8.5 scenario indicate a larger warming in the north-western part of the basin (*i.e.* in the Arabian Sea and western equatorial IO), a warming pattern that has sometimes been referred as “IOD-like”. The CMIP5 MMM and observations yield a similar, weaker amplitude, pSSTc over the 1871-2016 period, despite a stronger amplitude in observations, in particular for the relative cooling in the southern hemisphere. Our results indicate that part of this mismatch can be attributed to a strong aliasing of the pSSTc estimate by internal climate variability, even over the entire historical period (1871-2005). The effect of internal climate variability on the IO average SST is comparatively much smaller.

We further estimate emergence times (*i.e.* when the trend “signal” significantly emerges from the climate variability “noise”) for the NS, WE and AB indices, as well as for the IO-averaged warming. The IO-averaged warming signal is clearly detectable in all models, with a median emergence time around 1930. A warmer Arabian Sea than Bay of Bengal also emerges before 2020 in 80% of the models, indicating that this change is already detectable in CMIP5. Changes in the NS and WE gradients are on the other hand currently not detectable in most CMIP5 models. The warmer Northern than Southern tropical IO is indeed only detectable for 20% of the models before 2020, and the warmer western than eastern equatorial Indian Ocean for 40%.

Observations agree with CMIP5 models in indicating that the stronger warming in the western than in the eastern Indian ocean is not detectable by 2016. Unlike in CMIP5 models, the stronger Arabian sea warming is not detectable in observations by 2016, with observations at the lower end of the model distribution. Observations also contradict CMIP5 models in indicating a detectable stronger warming of the Northern than of the Southern IO. Subsampling the models and observations to only retain grid cells which are weakly influenced by representation error however strongly

reduces this mismatch. This suggests that the stronger northern IO warming in observations is overestimated, probably in relation with the poor data coverage in the southern hemisphere during the earlier part of the observational record. Overall, no SST gradient change is currently detectable in both the models and observations, due to a combination of aliasing by climate variability and observational uncertainties. In contrast, the IO-averaged warming is detectable since several decades in CMIP5 and observations, and only retaining the most accurate observations increases the observed warming, making it more consistent with the CMIP5 median.

5.2 Discussion: other observational products

Most of our observational analyses were so far obtained using HadSST3, which does not use any interpolation method to fill gaps in data-sparse regions. There are however other data products available over the historical period, such as HadiSST, ERSST-V3b, ERSST-V4, ERSST-V5, COBE-SST, COBE-SST2 and KAPLAN-SST. Contrary to HadSST3, those products rely on spatial interpolation, based on an empirical orthogonal function (EOF) analysis that extracts the main modes of variability of the data. Here, we will qualitatively compare observational results obtained with HadSST3 with those obtained from these products. Figure 10 displays the 1871-2016 SST trends for all the above products. (Note that the COBE products are only available from 12891 onward, and that the trend was hence estimated over this period for COBE. The SST trends of the other products are however very robust when estimated over 1891-2016 rather than over 1871-2016, indicating that the differences between the COBE and other products arise from the data processing, not from the trend computation period.) Figure 11 displays the SST trend for the IO average and the three SST gradient indices for all the above observational products. Additional analyses (not shown) indicate that the trends computed from interpolated SST products are marginally affected when subsampling these products to only retain the HAdSST3 valid data points. I.e. the differences between HAdSST3 and other products do not result from the sensitivity of the trend computation to missing values, but rather to other details in the data processing, such as the quality control, and the non-local nature of the data influence in interpolated products (i.e. the fact that a data at a given location can influence the trend at other locations).

All products indicate a similar average IO warming rate over the historical period (Fig. 11a), ranging between $\sim 0.7^{\circ}\text{C}$ and $\sim 0.8^{\circ}\text{C}$. Most datasets yield a qualitatively consistent basin-scale SST pattern (Fig. 10), with a maximum relative warming in the southern Arabian Sea, and a cooling (relative to the IO mean) in the southern IO. It is however not the case for the KAPLAN product, which is clearly an outlier (Fig. 10h), with a maximum warming in the central IO. Although the basin-scale structure of the trend pattern looks similar in the other products, there are non-negligible

amplitude and/or sign differences at a more regional scale (Fig. 10a-g). For instance, some products indicate statistically significant positive trends in the Mozambique channel, along the Java/Sumatra coast and Bay of Bengal, while some other rather point to negative trends. Similarly, there is a strong and localized warming trend along the ship tracks in the southern Arabian Sea and Bay of Bengal (cf. Fig. 8) in HadSST3 (Fig. 10a) and COBE-SST (Fig. 10f). This warming is considerably weaker in the southern Bay of Bengal and extends more into the Arabian Sea in ERSST datasets (Fig. 10 c, d and e). These regional differences result in non-negligible differences in the “SST gradients” trends (Fig. 11,b,c,d). There is for instance an insignificant WE gradient change in HadSST3 and COBE-SST, unlike in other products (Fig. 11c). Similarly, a weaker northern AS warming (Fig. 10) yields an insignificant positive change in the AB gradient in HadSST3 (Fig. 11b), unlike in other products (except KAPLAN). We think that these discrepancies between observational products can probably be attributed to the EOF projection technique used for spatial interpolation. Leading EOFs in the IO indeed capture a relatively homogeneous basin-scale pattern in response to remote forcing from the El Niño Southern Oscillation and an east-west dipole pattern associated with the IOD (e.g. Deser et al. 2010). These signals may have quite different spatial structures from that associated with anthropogenic forcing, and constraining these patterns with a warming in the western IO (where there is the maximum number of observations, Fig. 8) may result in an emphasized western warming in relation with the IOD pattern. This might explain the stronger WE change values in most observational products, relative to HadSST3 (Fig. 11c). For this reason, we think it is better to study long-term trends using a product that does not involve any spatial interpolation such as HadSST3. While such a product has gaps, and is noisier than spatially-interpolated product, the availability of representation error allows a subsampling of the model and data to retain only the most reliable observational points, as we did in the current manuscript. We think that this is the best approach for comparing the long-term trends in the models and observations.

5.3 Conclusions

The strong aliasing of the externally-forced SST gradient changes by internal climate variability over the full 1871-2016 historical period implies that a single model realization cannot be used for a robust estimation of these gradient changes. This also means that externally-forced pSSTc estimated from observations are heavily aliased by internal climate variability, even over the ~150 years-long observational period. We think that the broad agreement between the modelled and observed 1871-2016 “IOD-like” SST change patterns is thus partly coincidental. Climate change studies hence need to use the longest possible period (the full historical record) to estimate SST

trends, and somewhat take aliasing by internal climate variability into account through careful significance testing.

Observational estimates of the trend are themselves influenced by aliasing. But the observational issues highlighted in this paper, in particular before 1960 and in the southern and eastern equatorial Indian Ocean, further suggest that trend estimates from observations should be interpreted cautiously. For instance, our results reveal that the larger warming in the Northern than in the Southern hemisphere in observations is most likely overestimated, as this contrast is greatly reduced when excluding the most uncertain grid cells (Fig. 9d). The similar spatial pattern between the regions of strongest warming (Fig. 1d) and the density of observations during the earlier part of the record (Fig. 8a, b) is also somewhat suspicious, suggesting that the trend pattern is itself influenced by the observational spatial sampling.

The main question asked in the current study is whether we can confirm the CMIP5 MMM “IOD-like” SST pattern, with a maximum SST warming in the Arabian Sea and western equatorial Indian Ocean. A stronger warming in the western than in the eastern equatorial Indian would arguably lead to a weakening of the Indian Summer Monsoon (Kulkarni, 2012; Roxy et al. 2014; Roxy et al. 2015) and future increase of extreme positive IOD events (Cai et al. 2014). Such warming is currently neither detected in CMIP5 models nor in observations. The stronger warming of the Arabian Sea than of the Bay of Bengal has been related to an observed increase in the Arabian Sea and a decrease in the Bay of Bengal cyclone frequency in the recent period (Murakami et al. 2013; Murakami et al. 2017). Our results indicate that this change in the AS SST is already detectable in most models, but not in HadSST3 observations, with observations at the lower end of the model distribution. Finally, Saha et al. (2014) argued that a post-1950 warming of the southern IO could explain a weakening of the Indian monsoon. Models and observations to the contrary indicate a stronger warming of the Northern hemisphere, but this warming is currently not detectable in models and probably overestimated in observation due to sampling issues. Overall, the strong aliasing by internal climate variability and observational issues together suggest that the broad agreement between the CMIP5 MMM and observed SST change patterns over the 1871-2016 period is coincidental. While the regional climate projections of CMIP5 models and their impact discussed above may be correct, it is thus currently difficult to verify them based on direct observational evidence. The unabated emissions in the coming decades however yield a fast increase in the proportion of models in which those signals are detectable (Fig. 7). A sustained ocean observational network in the Indian Ocean (Hermes et al. 2019) will thus most likely allow us to confirm or infirm CMIP5 projections. Already, the comparison with collocated observations suggest that the CMIP5

MMM overestimates the enhanced warming in the Arabian Sea relative to the Bay of Bengal (Fig. 9b).

Acknowledgements

This study was funded by the Council of Scientific and Industrial Research (CSIR), New Delhi. The authors thank the Director, CSIR-National Institute of Oceanography (CSIR-NIO), for his support. This work was done while T. I. and M. L. were visiting scientists at CSIR-NIO under Institute de Recherche pour le Development (IRD) funding. J. V. acknowledges IRD for supporting regular visits to CSIR-NIO and thanks CSIR-NIO for his “Adjunct Scientist” position. We acknowledge the World Climate Research Program’s Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1) for producing and making available their model outputs. All the CMIP5 model outputs were downloaded from the Institute Pierre Simon Laplace (IPSL) CICALAD repository. The NOAA FERRET software was used here for analysis. This is NIO contribution XXXX.

References

- Alory G, Meyers G (2009) Warming of the upper equatorial Indian Ocean and changes in the heat budget (1960–1999). *J Clim* 22:93–113. DOI: <https://doi.org/10.1175/2008JCLI2330.1>
- Alory G, Wijffels S, Meyers M (2007) Observed temperature trends in the Indian Ocean over 1960–1999 and associated mechanisms. *Geophys Res Lett* 34(2):L02606. DOI: <https://doi.org/10.1029/2006GL028044>
- Bretherton CS, Widmann M, Dymnikov VP, Wallace JM, Bladé I (1999) The effective number of spatial degrees of freedom of a time-varying field. *J Clim* 12(7):1990–2009. DOI: [https://doi.org/10.1175/1520-0442\(1999\)012<1990:TENOSD>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2)
- Cai, W., Santoso, A., Wang, G., Weller, E., Wu, L., Ashok, K., Masumoto, Y and Yamagata, T. (2014). Increased frequency of extreme Indian Ocean Dipole events due to greenhouse warming. *Nature*, Volume 510, p. 254-258. DOI: 10.1038/nature13327
- Deser C, Alexander MA, Xie SP, Phillips AS (2010) Sea surface temperature variability: Patterns and mechanisms. *Annual review of marine science* 2:115-143. DOI: 10.1146/annurev-marine-120408-151453
- Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2012). Uncertainty in climate change projections: the role of internal variability. *Climate dynamics*, 38(3-4), 527-546. DOI: <https://doi.org/10.1175/JCLI-D-12-00788.1>
- Dong, L., Zhou, T. & Wu, B. (2014) Indian Ocean warming during 1958–2004 simulated by a climate system model and its mechanism. *Clim. Dyn.* 42, 203–217. DOI: <https://doi.org/10.1007/s00382-013-1722-z>

- Dong L., and Tianjun Zhou. (2014). The Indian Ocean Sea Surface Temperature Warming Simulated by CMIP5 Models during the 20th Century: Competing Forcing Roles of GHGs and Anthropogenic Aerosols. *J. of Clim.*, 27, 3348-3362. <https://doi.org/10.1175/JCLI-D-13-00396.1>.
- Du, Y., and S.-P. Xie (2008), Role of atmospheric adjustments in the tropical Indian Ocean warming during the 20th century in climate models, *Geophys. Res. Lett.*, 35, L08712, DOI:10.1029/2008GL033631.
- Folland, C. K. and D. E. Parker. (1995). Correction of instrumental biases in historical sea surface temperature data. *Q. J. R. Meteorol. Soc.*, 121, 319-367. <https://doi.org/10.1002/qj.49712152206>
- Funk CC, Dettinger MD, Michaelsen JC, Verdin JP, Brown ME, Barlow M, Hoell A (2008) Warming of the Indian Ocean threatens eastern and southern African food security but could be mitigated by agricultural development. *Proc Nat Acad Sci USA*, 105:11081-11086. DOI: 10.1073/pnas.0708196105
- Hamlington B D, Strassburg M W, Leben R R, Han W, Nerem R S and Kim K-Y (2014) Uncovering an anthropogenic sea-level rise signal in the Pacific Ocean *Nat. Clim. Change* 4 782–5. DOI: <https://doi.org/10.1038/nclimate2307>
- Han, W., J. Vialard, M.J. McPhaden, T. Lee, Y. Masumoto, M. Feng and W. de Ruijter, (2014) Indian Ocean Decadal Variability: A Review, *Bull. Am. Met. Soc.*, 95, 1679-1703. DOI: <https://doi.org/10.1175/BAMS-D-13-00028.1>
- Hawkins E, Sutton R (2012) Time of emergence of climate signals. *Geophys Res Lett* 39:n/a–n/a. DOI: 10.1029/2011GL050087
- Hegerl GC, Karl TR, Allen M, Bindoff NL, Gillett N, Karoly D, Zhang X, Zwiers F (2006) Climate change detection and attribution: beyond mean temperature signals. *J Clim* 19(20):5058–5077. DOI: 10.1175/JCLI3900.1
- Hermes, J., et al.(2019).A sustained ocean observing system in the Indian Ocean for climate related scientific knowledge and societal needs.*Front. Mar. Sci.* DOI : <https://doi.org/10.3389/fmars.2019.00355>
- Hirahara, S., Ishii, M., and Y. Fukuda, (2014) Centennial-scale sea surface temperature analysis and its uncertainty. *J of Climate*, 27, 57-75. <https://doi.org/10.1175/JCLI-D-12-00837.1>
- Hoerling, M., Eischeid, J., Perlwitz, J., Quan, X., Zhang, T. and Pegion, P. (2012) On the increased frequency of Mediterranean drought. *J. Climate* 25, 2146–2161. <https://journals.ametsoc.org/doi/pdf/10.1175/JCLI-D-11-00296.1>
- Huang, B., P. Thorne, T. Smith, W. Liu, J. Lawrimore, V. Banzon, H. Zhang, T. Peterson, and M. Menne, (2015) Further Exploring and Quantifying Uncertainties for Extended Reconstructed Sea Surface Temperature (ERSST) Version 4 (v4). *Journal of Climate*, 29, 3119–3142, DOI:10.1175/JCLI-D-15-0430.
- Huang, B., Peter W. Thorne, et. al, (2017). Extended Reconstructed Sea Surface Temperature version 5 (ERSSTv5), Upgrades, validations, and intercomparisons. *J. Climate*, doi: 10.1175/JCLI-D-16-0836.1

- Ihara, C., Y. Kushnir, and M. A. Cane, (2008) Warming trend of the Indian Ocean SST and Indian Ocean dipole from 1880 to 2004. *J. Climate*, 21, 2035–2046, DOI: 10.1175/2007 JCLI1945.1.
- Ihara, C., Kushnir, Y., Cane, M. A., & de La Peña, V. H. (2009). Climate change over the equatorial Indo-Pacific in global warming. *Journal of Climate*, 22(10), 2678–2693. DOI:10.1175/2008jcli2581.1
- Izumo, T., Lengaigne, M., Vialard, J. et al. (2014). Influence of Indian Ocean Dipole and Pacific recharge on following year's El Niño: interdecadal robustness. *ClimDyn.* 42: 291. <https://journals.ametsoc.org/doi/pdf/10.1175/JCLI-D-15-0481.1>
- Kaplan, A., M. Cane, Y. Kushnir, A. Clement, M. Blumenthal, and B. Rajagopalan. (1998). Analyses of global sea surface temperature 1856-1991, *Journal of Geophysical Research*, 103, 18,567–18,589.
- Kennedy J.J., Rayner, N.A., Smith, R.O., Saunby, M. and Parker, D.E. (2011a). Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 1: measurement and sampling errors. *J. Geophys. Res.*, 116, D14103, DOI:10.1029/2010JD015218.
- Kennedy J.J., Rayner, N.A., Smith, R.O., Saunby, M. and Parker, D.E. (2011b). Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 2: biases and homogenisation. *J. Geophys. Res.*, 116, D14104, DOI:10.1029/2010JD015220
- Kulkarni, A., (2012) Weakening of Indian summer monsoon rainfall in warming environment. *Theor. Appl. Climatol.*, 109(3–4), 447–459, DOI: 10.1007/s00704-012-0591-4.
- Levitus, S., J. I. Antonov, T. P. Boyer, and C. Stephens, (2009) Global ocean heat content 1955-2008 in light of recently revealed instrumentation problems. *Geophys. Res. Lett.*, 36, L07608, DOI:10.1029/2008GL037155.
- Liu, W., Lu, J. and Xie. S-P., (2015) Understanding the Indian Ocean response to double CO₂ forcing in a coupled model. *Ocean Dynamics* 65:1037-1046. DOI:10.1007/s10236-015-0854-6
- Luo, J.-J., W. Sasaki, and Y. Masumoto (2012), Indian Ocean warming modulates Pacific climate changes, *Proc. Natl. Acad. Sci. U.S.A.*, 109(18) 701–18,706. DOI: 10.1073/pnas.1210239109
- Murakami H, Sugi M, Kitoh A (2013) Future changes in tropical cyclone activity in the North Indian Ocean projected by high-resolution MRI-AGCMs. *ClimDyn* 40:1949-1968. DOI:<https://doi.org/10.1007/s00382-012-1407-z>.
- Murakami, Hiroyuki & Vecchi, Gabriel & Underwood, Seth. (2017). Increasing frequency of extremely severe cyclonic storms over the Arabian Sea. *Nature Climate Change*. 7. DOI:10.1038/s41558-017-0008-6.
- Murtugudde, R., J. P. McCreary, and A. J. (2000). Busalacchi, Oceanic processes associated with anomalous events in Indian Ocean with relevance to 1997-1998, *J. Geophys. Res.*, 105(3295–3306). DOI: <https://doi.org/10.1029/1999JC900294>
- Parvathi, V., Suresh, I., Lengaigne, M., Izumo, T., & Vialard, J. (2017). Robust projected weakening of winter monsoon winds over the Arabian Sea under climate change. *Geophysical Research Letters*, 44. DOI:<https://doi.org/10.1002/2017GL075098>.

- Rahul S, Gnanaseelan C (2016) Can large scale surface circulation changes modulate the sea surface warming pattern in the Tropical Indian Ocean?. *Climate Dynamics*, 46: 3617-3632, DOI:10.1007/s00382-015-2790-z
- Rayner, N. A.; Parker, D. E.; Horton, E. B.; Folland, C. K.; Alexander, L. V.; Rowell, D. P.; Kent, E. C.; Kaplan, A. (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century *J. Geophys. Res.* Vol. 108, No. D14, 4407 DOI:10.1029/2002JD002670.
- Reverdin, G., D. Cadet, and G. D.(1986). Interannual displacements of convection and surface circulation over the equatorial Indian Ocean, *Quart. J. Roy. Meteorol. Soc.*,122(43–67).<https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.49711247104>
- Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., ... &Rafaj, P. (2011). RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change*, 109(1-2), 33.DOI:10.1007/s10584-011-0149-y.
- Rogelj, J., Elzen, M., Höhne, N., Fransen, T., Fekete, H., Winkler, H., Schaeffer, R., Sha, F., Riahi, K., and Meinshausen, M. (2016), Paris Agreement climate proposals need a boost to keep warming well below 2 °C, *Nature*, 534, 631– 639.DOI: 10.1038/nature18307
- Roxy M. K., K. Ritika, P. Terray, S. Masson (2014). The curious case of Indian Ocean warming. *J.Climate*, 27, 22, 8501-8509.DOI:10.1175/JCLI-D-14-00471.1
- Roxy, M. K., K. Ritika, P. Terray, R. Murtugudde, K. Ashok, and B. N. Goswami (2015), Drying of Indian subcontinent by rapid Indian Ocean warming and a weakening land-sea thermal gradient, *Nat. Commun.*, 6, 7423.DOI: 10.1038/ncomms8423
- Saha, A., S. Ghosh, A. S. Sahana, and E. P. Rao (2014), Failure of CMIP5 climate models in simulating post-1950 decreasing trend of Indian monsoon, *Geophys. Res. Lett.*, 41, 7323–7330, doi:10.1002/2014GL061573.
- Saji, N. H., Goswami, B. N., Vinayachandran, P. N. and Yamagata, T.(1999), A dipole mode in the tropical Indian Ocean. *Nature* 401, 360-363.DOI: 10.1038/43854.
- Smith, T.M., R.W. Reynolds, T.C. Peterson, and J. Lawrimore.(2008): Improvements NOAAs Historical Merged Land–Ocean Temp Analysis (1880–2006). *Journal of Climate*, **21**, 2283–2296. DOI: 10.1175/2007JCLI2100.1
- Swapna, P., Krishnan, R., & Wallace, J. M. (2014). Indian Ocean and monsoon coupled interactions in a warming environment. *Climate Dynamics*, 42(9–10), 2439–2454. <https://doi.org/10.1007/s00382-013-1787-8>.
- Taylor KE, Stouffer BJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *B Am Meteorol Soc* 93:485–498. <https://doi.org/10.1175/BAMS-D-11-00094.1>
- Terray, P., (1994) An evaluation of climatological data in the Indian Ocean area. *J. Meteor. Soc. Japan*, 72, 359–386.DOI: 10.1038/nature11576
- Tokinaga, H., Xie, S.-P., Deser, C., Kosaka, Y., &Okumura, Y.-M. (2012). Slowdown of the Walker circulation driven by tropical Indo-Pacific warming. *Nature*, 491, 439–443.DOI: 10.1038/nature11576

- Vecchi, G. A. & Soden, B. J. Global warming and the weakening of the tropical circulation. *J. Clim.* 20, 4316–4340 (2007). DOI: <https://doi.org/10.1175/JCLI4258.1>
- Webster, P. J., Moore, A., Loschnigg, J., and Leban, M.: 1999, Coupled dynamics in the Indian Ocean during 1997-1998. *Nature*, 401, 356-360. DOI: <https://doi.org/10.1038/43848>
- Woodruff, S., S. Worley, S. Lubker, Z. Ji, J. Freeman, D. Berry, P. Brohan, E. Kent, R. Reynolds, S. Smith, and C. Wilkinson (2010), ICOADS release 2.5: extensions and enhancements to the surface marine meteorological archive, *International Journal of Climatology*, DOI: <https://doi.org/10.1002/joc.2103>.
- Yao, SL., Huang, G., Wu, RG. et al. *ClimDyn* and Chen D. (2016). Inhomogeneous warming of the Tropical Indian Ocean in the CMIP5 model simulations during 1900–2005 and associated mechanisms. *ClimDyn* 46:619–636. doi: <https://doi.org/10.1007/s00382-015-2602-5>
- Yue, S & Pilon, P. (2004) A comparison of the power of the t test, Mann-Kendall and bootstrap tests for trend detection / Une comparaison de la puissance des tests t de Student, de Mann-Kendall et du bootstrap pour la détection de tendance, *Hydrol. Sci.* 49:1(21-37). DOI: <https://doi.org/10.1623/hysj.49.1.21.53996>
- Zheng, X. T., Xie, S. P., Du, Y., Liu, L., Huang, G., & Liu, Q. (2013). Indian Ocean Dipole response to global warming in the CMIP5 multimodel ensemble. *J. Climate*, 26(16), 6067–6080. DOI: <https://doi.org/10.1175/JCLI-D-12-00638>.

Institute	Model	Model # for
BCC	bcc-csm1-1	15
	bcc-csm1-1-m	5
BNU	BNU-ESM	8
CCCma	CanESM2	25
CMCC	CMCC-CESM	17
	CMCC-CMS	30
CNRM-	CNRM-CM5	28
CSIRO-BOM	ACCESS1-3	21
CSIRO-OCCCE	CSIRO-Mk3-6-0	23
FIO	FIO-ESM	4
IPSL	IPSL-CM5A-LR	7
	IPSL-CM5A-MR	18
	IPSL-CM5B-LR	27
LASG-CESS	FGOALS-g2	36
LASG-IAP	FGOALS-s2	16
MIROC	MIROC5	32
	MIROC-ESM	33
	MIROC-ESM-	34
MOHC	HadGEM2-CC	14
MPI-M	MPI-ESM-LR	6

	MPI-ESM-MR	3
MRI	MRI-CGCM3	10
	MRI-ESM1	13
NASA-GISS	GISS-E2-H	20
	GISS-E2-H-CC	26
	GISS-E2-R	31
	GISS-E2-R-CC	29
NCAR	CCSM4	12
NCC	NorESM1-M	19
	NorESM1-ME	35
NIMR-KMA	HadGEM2-AO	9
NOAA-GFDL	GFDL-CM3	24
	GFDL-ESM2G	22
	GFDL-ESM2M	1
NSF-DOE-NCAR	CESM1-BGC	11
	CESM1-CAM5	2

Table 1. List of the 36 CMIP5 models used in this study (in alphabetic order, based on the institute name) and the number by which each model is designated on Figure 5 (see text and figure captions for more details).

	Emergence time vs long-term index change (signal)	Emergence time vs index internal variability (noise)
IO	-0.37	0.57
AB	-0.51	0.39
WE	-0.77	0.46
NS	-0.50	0.41

Table 2. The inter-model partial correlation between the emergence time and the signal / noise amplitude, for various surface temperature indices (IO: Indian Ocean average, AB: Arabian Sea minus Bay of Bengal). See text for details on the partial correlation analysis and how the signal (long-term change in the index) and noise (amplitude of year-to-year to decadal variability in the index) are estimated. All the correlations in this table are significantly different from zero at the 95% level.

%	AS	BOB	NIO	SIO	WEST	EAST
HAD3	7	67	7	6	7	9
SUB 0.5	5	52	5	2	4	6
SUB 0.4	5	45	4	1	2	5
SUB 0.3	4	32	3	4	1	3

Table 3.Percentage of HadSST3 valid grid points (GP)in the Arabian Sea (AS), Bay of Bengal (BoB), Northern Indian Ocean (NIO), Southern tropical Indian Ocean (SIO), western (WEIO) and Eastern (EEIO) equatorial Indian Ocean, for the HadSST3 observations and when subsampling them based on three observational error thresholds (0.5°C, 0.4°C and 0.3°C, see text for details).

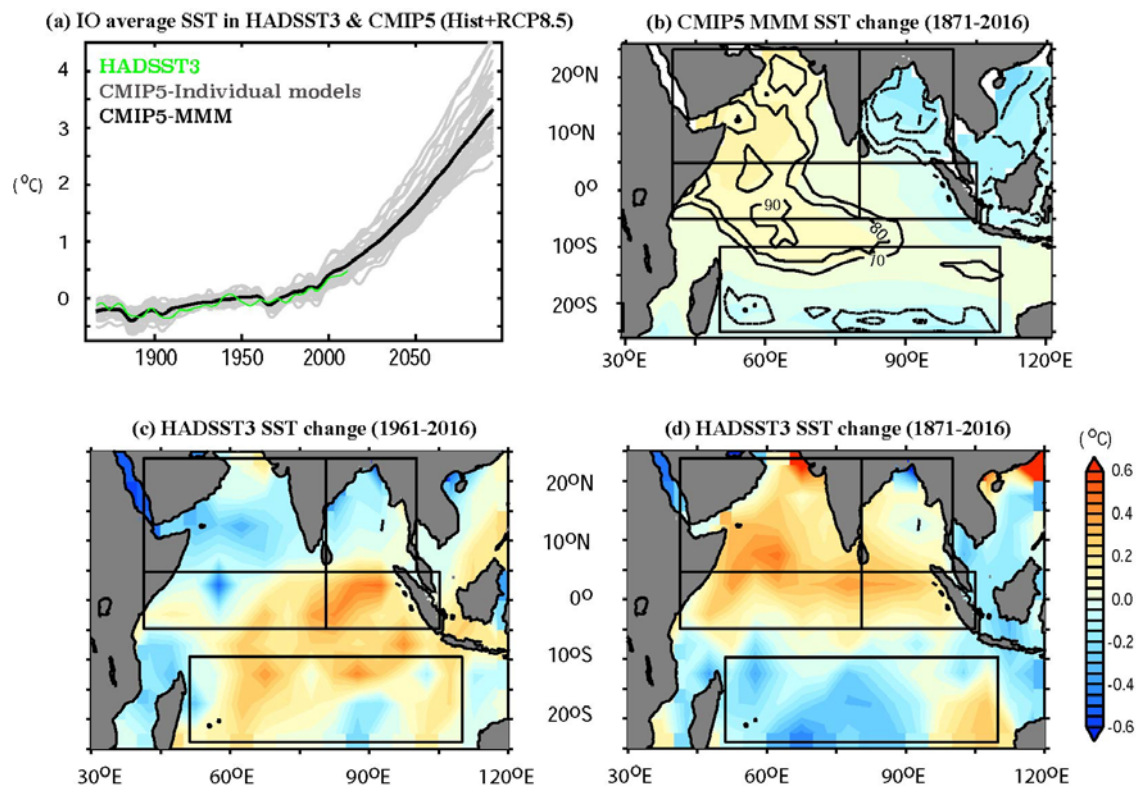


Figure 1. Indian Ocean (IO) SST trend in CMIP5 models and observations. (a) Time evolution of the average IO (north of 25°S) SST anomalies (°C; relative to the 1871-2016 climatology), for the HADSST3 dataset (green curve), the 36 CMIP5 models historical and RCP8.5 simulations (grey curves) and the multi-model mean (thick black curve). (b) Map of the CMIP5 multi-model mean SST change (°C) relative to the IO mean change over the 1871-2016 period (linear trend estimate, see text for details), with black continuous (dashed) contours indicating when the local warming is larger (weaker) than the tropical IO mean warming for at least 70%, 80% and 90% of the models. (c, d) Same as (b), but for HADSST3 observations over the (c) recent 1961-2016 period and (d) the extended 1871-2016 historical period. Black frames on panels b, c and d delineate boxes used for defining the indices in this paper (see section 2 for details).

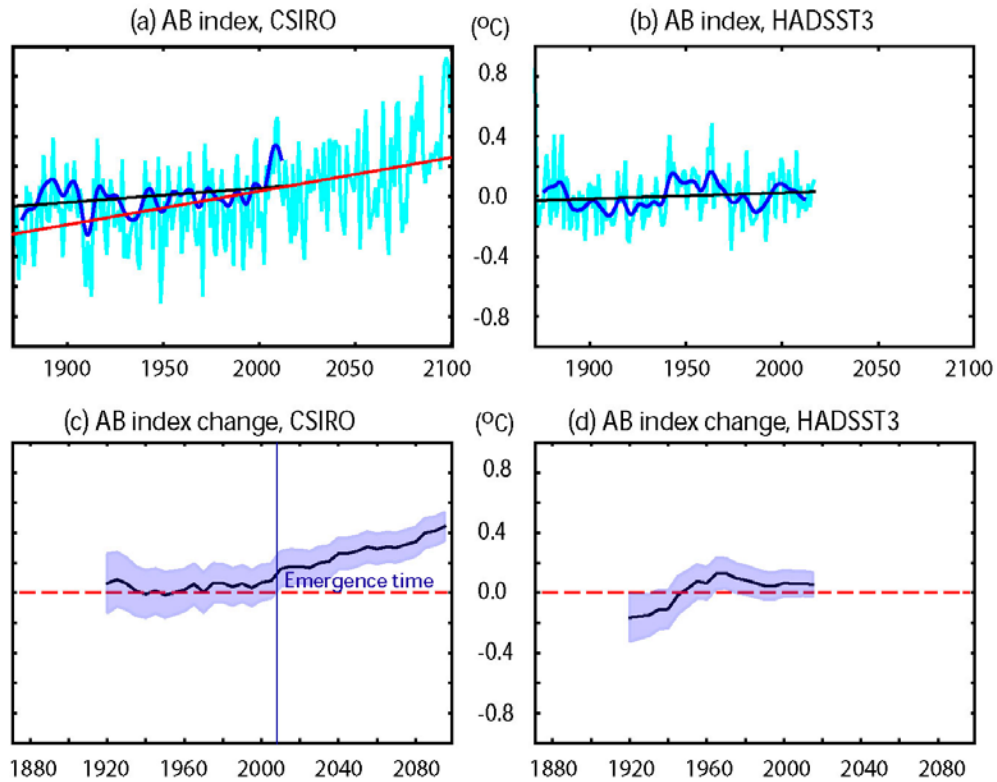


Figure 2. Method for estimating emergence time. (a, b) Time series of Arabian Sea minus Bay of Bengal (AB) SST anomalies (relative to the 1871-2016 climatology) for the annual-mean data (light blue curves) and 12-year Hanning window low-pass-filtered data (dark blue curves) until 2016 for (a) a given model (CSIRO-MK3) and (b) HadSST3 dataset. The linear trend estimated over 1871-2016 is indicated as a black line in panels a and b, and that over 1871-2100 as a red line in panel a. (c, d) Time series of AB SST change (black curves, obtained from the trend multiplied by the length of the period, cf. section 2) estimated over a period starting in 1871 with an end point varying between 1920 and 2100 (2016 for observations), with the blue shading on panels c,d indicating the 90% confidence interval. The vertical blue line in panel c indicates the emergence time.

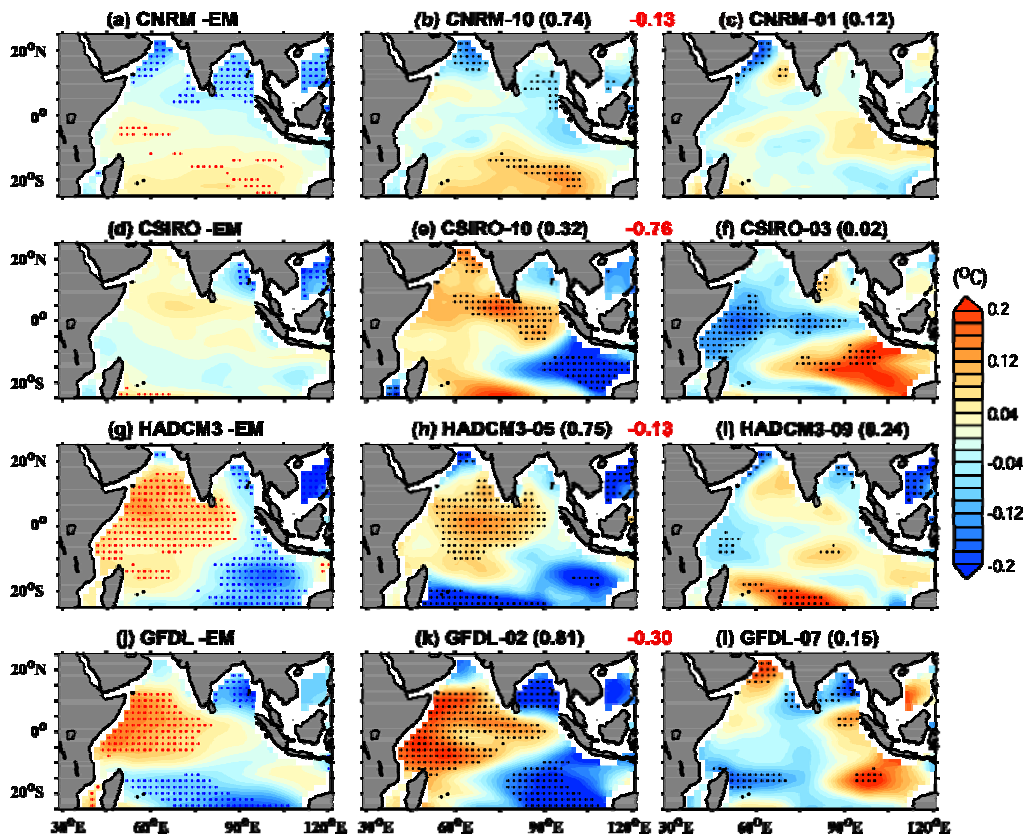


Figure 3. Spatial pattern of IO relative SST change for the ensemble mean and selected members in four ensemble experiments ($^{\circ}\text{C}$, relative to the basin-averaged trend) over the 1871-2005 historical period. (Left column) 10-member ensemble mean SST trend of (a) CNRM-CM5, (d) CSIRO, (g) HadCM3 and (j) GFDL models, with red (blue) dots indicating where the local warming is larger (weaker) than the tropical IO-averaged warming for 90 % or more of the ensemble members. (Middle and right columns) SST trend for the two members with the weakest pattern correlation (indicated in red above the panels) for (b,c) CNRM-CM5, (e,f) CSIRO-MK3-6, (h,i) HadCM3 and (k,l) GFDL2.1 models. The number within brackets above in the title of each middle and right column panel indicates the pattern correlation between the SST change of that member and the ensemble mean. Black dots on the middle and right columns indicate when this trend is significantly different from zero at the 90% level.

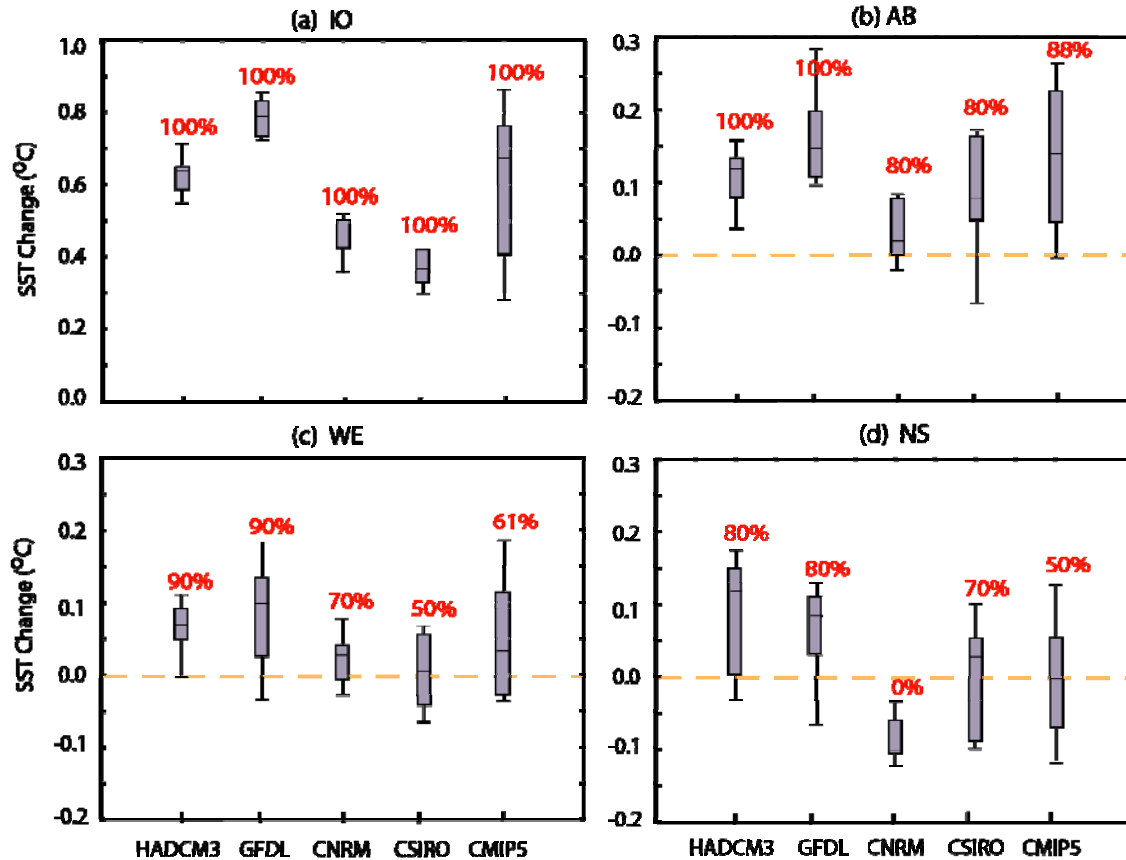


Figure 4. Model distribution of trends for four SST indices (°C). Median, first and last quartiles (boxes), and first and last deciles (whiskers) of the 1871-2005 SST change for (a) IO (Indian Ocean average), (b) AB (Arabian Sea minus Bay of Bengal), (c) WE (Western minus Eastern Equatorial Indian Ocean) and (d) NS (Northern minus Southern Indian Ocean) within each of the four 10-member ensemble experiments from figure 3, and for the 36 CMIP5 model database (listed in Table 1). The percentage of members/models with a positive index change is indicated in red above each whisker. Note that the vertical scale is ~3 times smaller for panels b,c,d than for panel a (i.e. changes in gradients are much smaller than changes in the IO-mean SST).

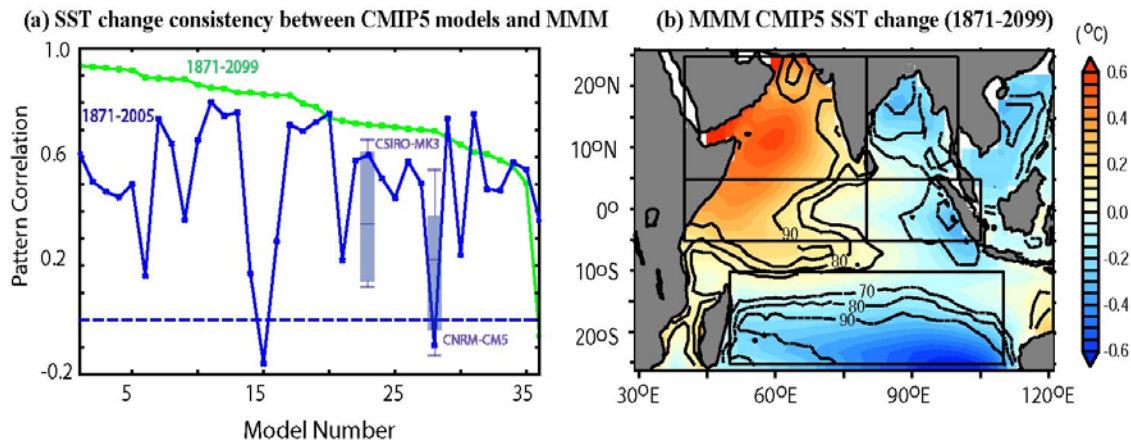


Figure 5. Inter-model agreement in the projected pattern of IO relative SST change. (a) pattern correlation between the CMIP5 MMM IO SST change and that of each individual model over the 1871-2099 (green) and 1871-2005 (blue) periods. The models are ranked in descending order based on the 1871-2099 pattern correlation (see Table 1 for correspondence between models and their numbers on the abscissa). The whiskers (median, first and last quartiles, and first and last deciles) indicate the distribution of 1871-2005 pattern correlation values for the two CMIP5 models with 10-member ensembles discussed in section 3.1 (#23 for CSIRO-Mk3-6-0 and #28 for CNRM-CM5). (b) Map of the CMIP5 MMM SST change ($^{\circ}\text{C}$) relative to the IO mean over the 1871-2099 period (i.e. as Fig. 1b, but for the historical+RCP8.5 simulations).

Annual time series of gradients indices in CMIP

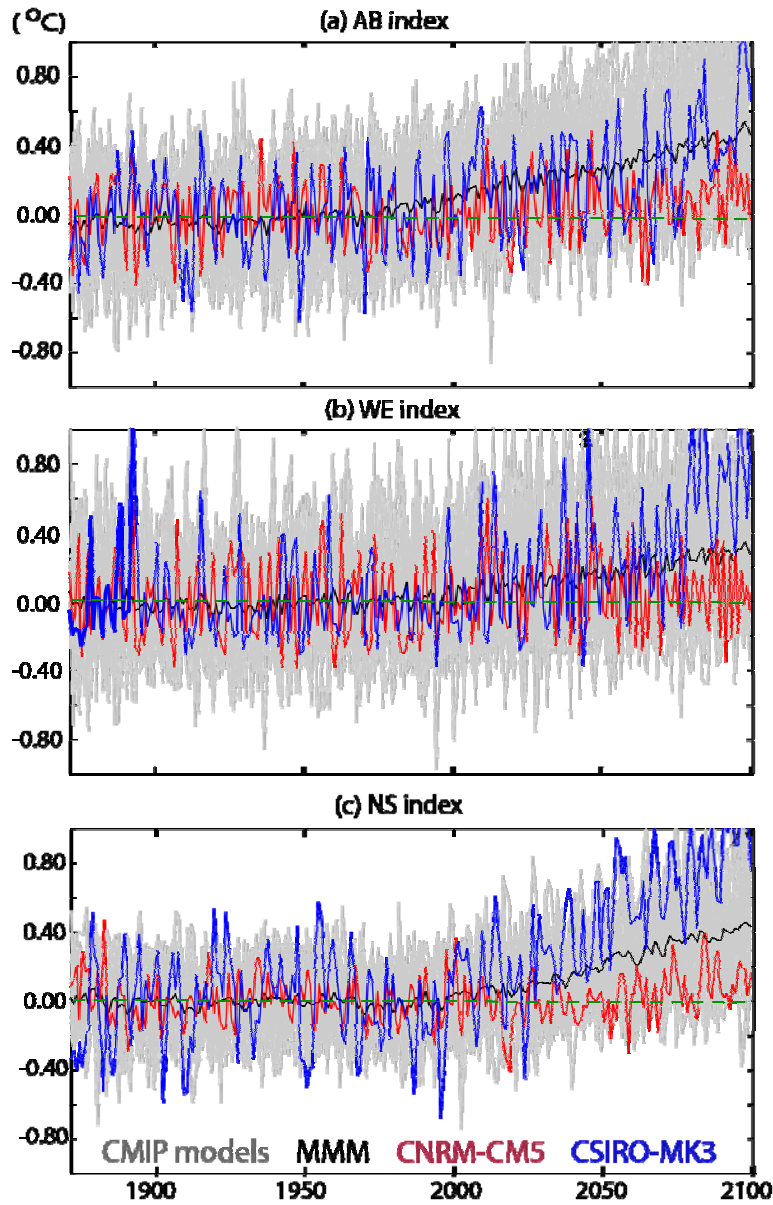


Figure6. Time series of annual-mean values of the SST gradient indices for the 36 CMIP5 models (grey lines), their MMM (black thick line) and two selected models (CNRM-CM5 in red, CSIRO-MK3 in blue) for (a) AB (Arabian Sea minus Bay of Bengal), (b) WE (West minus East equatorial Indian Ocean) and (c) NS (North minus South Indian Ocean) SST anomaly (relative to the 1871-2016 mean seasonal cycle).

Emergence time for 4 SST indices

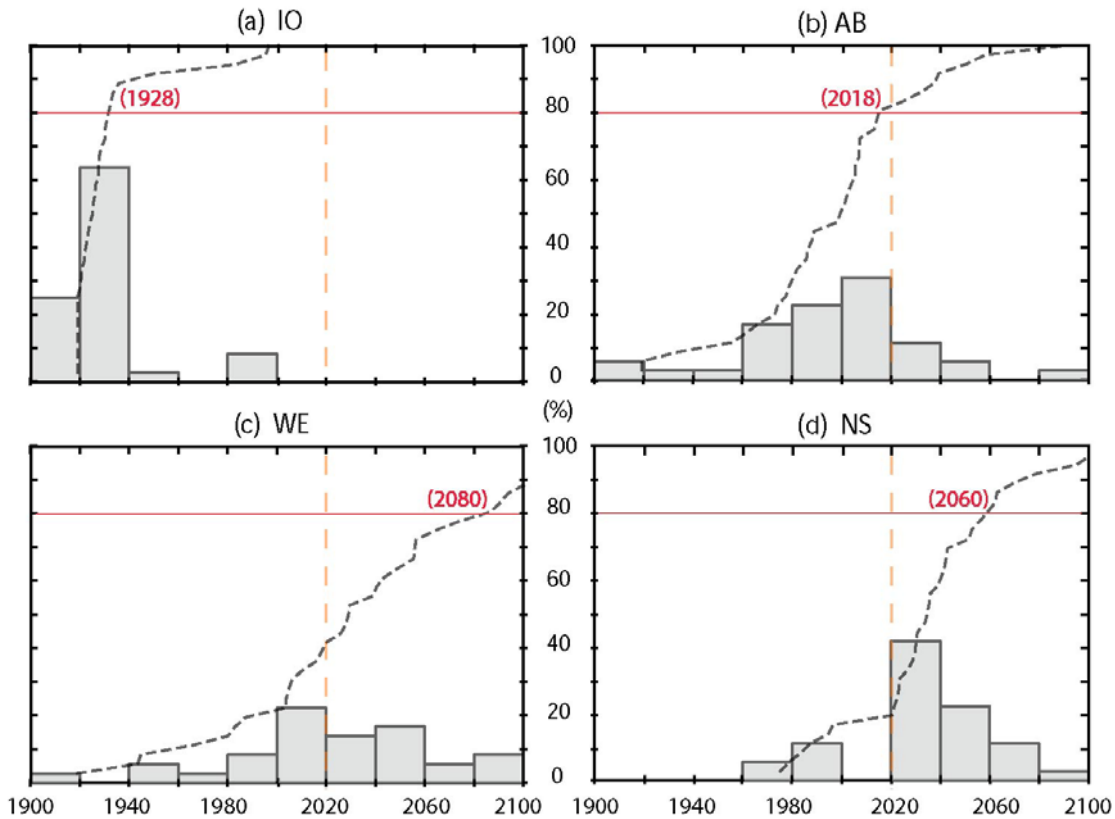


Figure 7. Emergence time distribution for four IO indices: Percentage of CMIP5 models with an emergence time within 20-year bins for (a) the IO-averaged warming, (b) a positive AB (i.e. larger Arabian Sea than Bay of Bengal warming), (c) a positive WE (i.e. larger Western than Eastern equatorial Indian Ocean warming) and (d) a positive NS (i.e. larger Northern than Southern tropical Indian Ocean warming). The dashed black curve shows the cumulative distribution of the modelled emergence time. The vertical dashed orange line marks 2020 to symbolize “present”. The red thin line indicates 80%: we consider that there is a consensus about the emergence of a signal in the CMIP5 database when $\geq 80\%$ of the models indicate a detectable signal. The year when changes in various indices become detectable in CMIP5, according to this definition, is indicated on each panel.

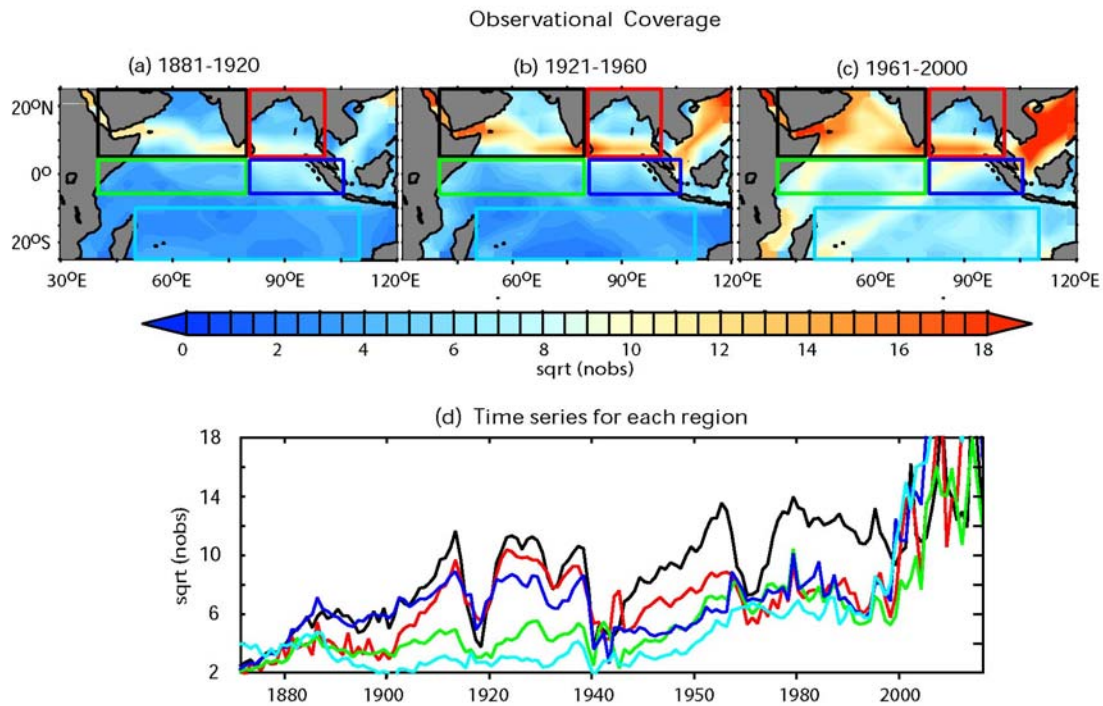


Figure 8. Spatio-temporal distribution of the number of surface temperature observations. Average maps of the square-root of the number of observations used in the HadSST3 gridded SST product per $5^\circ \times 5^\circ \times 1$ month for the (a) 1881-1920, (b) 1921-1960 and (c) 1961-2000 periods (as the sampling uncertainty depends of this square-root). (d) Corresponding time-series of the spatial average of this square-root within each of the boxes used to compute the SST gradient indices, overlaid in colour on panels a,b,c: Arabian Sea (black); Bay of Bengal (red); Western (green) and Eastern (blue) equatorial Indian Ocean; Southern tropical Indian Ocean (cyan).

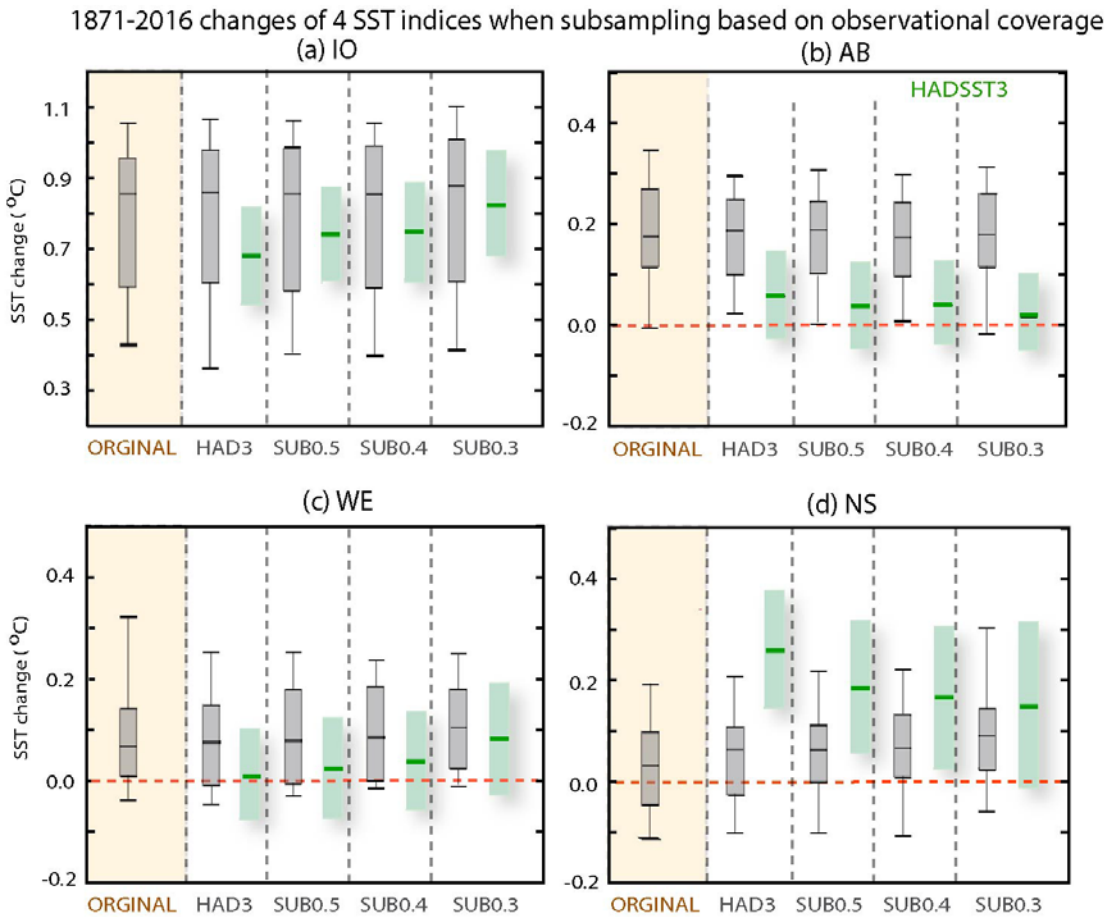


Figure 9. Influence of spatial sampling on the SST change distribution of four IO SST indices. CMIP5 model distribution (the box indicates the median, 1st and last quartiles, and whiskers the 1st and last deciles) and HadSST3 observed value (green) of the 1871-2016 changes in various SST indices, with an increasingly severe (left to right) level of subsampling of the CMIP5 models and HadSST3, to retain only the most accurate HadSST3 grid points for: **(a)** IO (Indian Ocean average), **(b)** AB (Arabian Sea minus Bay of Bengal), **(c)** WE (Western minus Eastern equatorial Indian Ocean), and **(d)** NS (Northern minus Southern tropical Indian Ocean) SST change indices. **ORIGINAL:** Model without subsampling. **HAD3:** HadSST3 data coverage, with the model subsampled as HadSST3. **SUB0.5, SUB0.4, SUB0.3:** Model and HadSST3 subsampled to retain only grid points with an HadSST3 observational error estimate below 0.5, 0.4 and 0.3°C, respectively. The green shading represents the 90% confidence interval for HadSST3.

Observed estimates of 1871-2016 SST changes

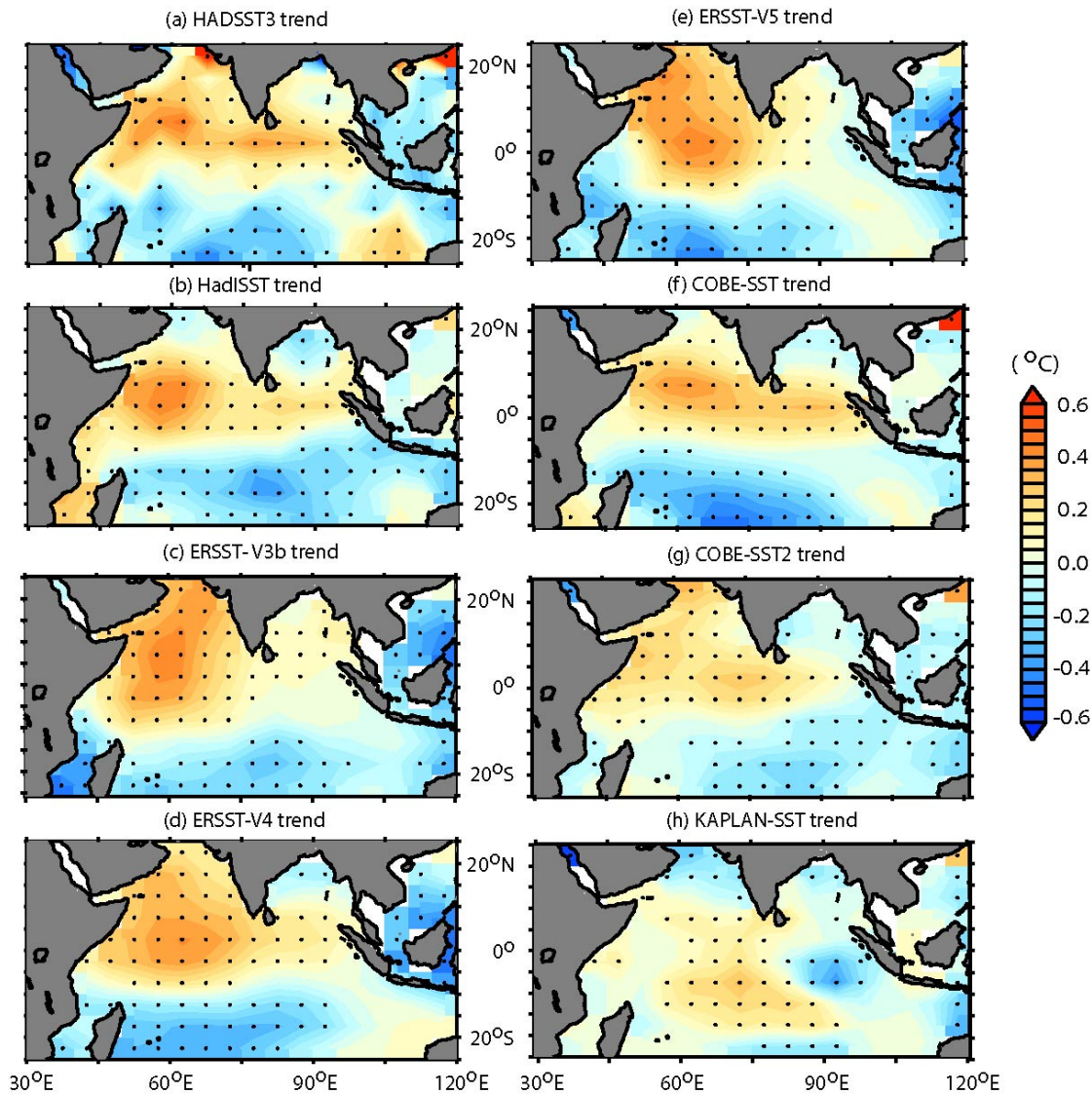


Figure 10. Estimates of the SST trends in observational products. 1871-2016 relative SST changes estimates from (a) HADSST3, (b) HadISST, (c) ERSSTV3b, (d) ERSST-V4, (e) ERSST-V5, (f) COBE-SST, (g) COBE-SST2, (h) KAPLAN-SST. Black dots indicate regions where the SST change is significantly different from zero at the 90% confidence level. Note that the COBE SST products are only available from 1891 to 2016, and that their SST change was thus estimated over this period. The SST trends of the other products are however very robust when estimated over 1891-2016 rather than over 1871-2016, indicating that the differences between the COBE products and the other products arise from the data processing, not from the trend computation period.

1871-2016 changes of 4 SST indices for all th observational products

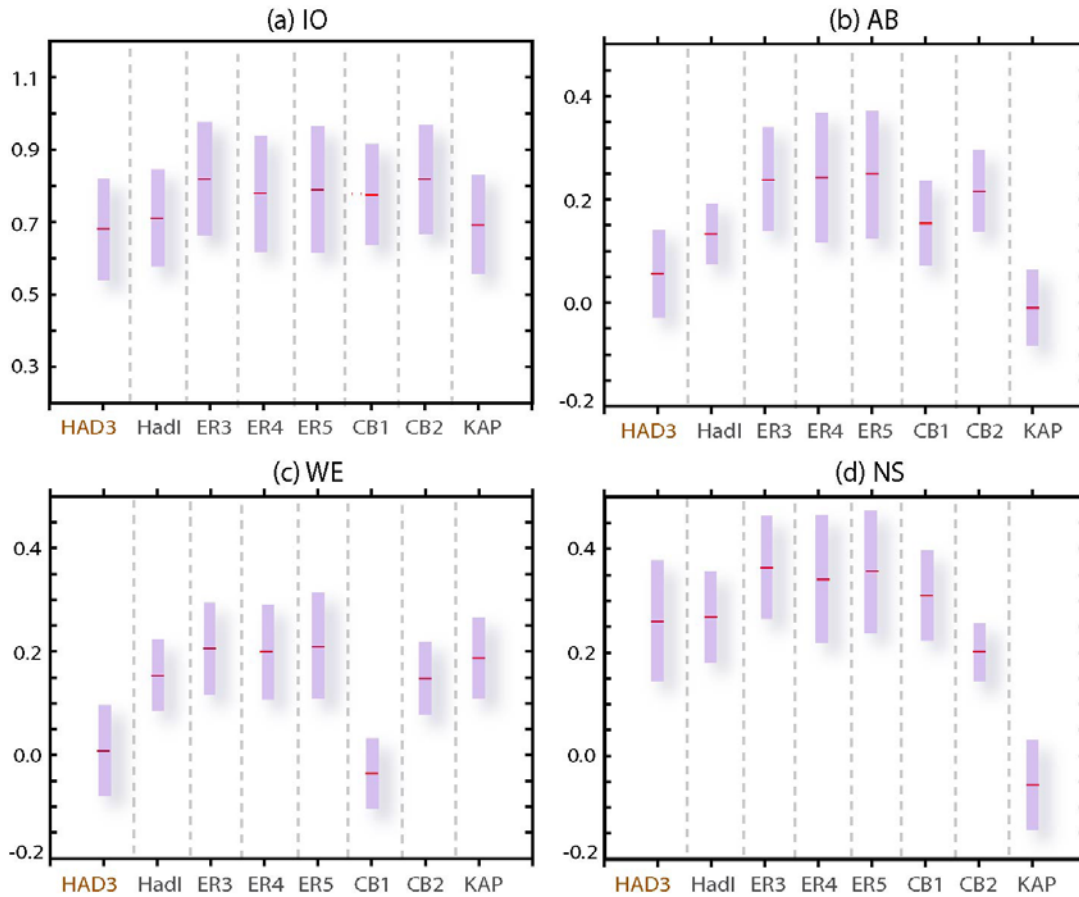


Figure 11. Estimates of the 1871-2016 SST changes for the (a) IO (Indian Ocean average), (b) AB (Arabian Sea minus Bay of Bengal), (c) WE (Western minus Eastern equatorial Indian Ocean), and (d) NS (Northern minus Southern tropical Indian Ocean) indices, from eight different observational products: HadSST3 (HAD3), HadISST (HadI), ERSST V3b, V4 and V5 (ER3, ER4 and ER5), COBE-SST and COBE-SST2 (CB1 and CB2) and KAPLAN (KAP). The shading indicates the 90% confidence interval on the trend estimate. Note that the COBE SST products are only available from 1891 to 2016, and that their SST change was thus estimated over this period. The SST trends of the other products are however very robust when estimated over 1891-2016 rather than over 1871-2016, indicating that the differences between the COBE products and the other products arise from the data processing, not from the trend computation period.