The Emerging Human Influence on the Seasonal Cycle of Sea Surface Temperature

Jia-Rui Shi[1*], Benjamin D. Santer[1,2*], Young-Oh Kwon[1], Susan E. Wijffels[1]

[1]*Woods Hole Oceanographic Institution, Woods Hole, MA, USA*

[2]*Joint Institute for Regional Earth System Science and Engineering, University of California Los Angeles, Los Angeles, CA, USA*

**Corresponding authors*:

Benjamin D. Santer, bensanter1289@gmail.com

Jia-Rui Shi, jia-rui.shi@whoi.edu

Physical Oceanography Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543.

Abstract

We provide the first scientific evidence that a human-caused signal in the seasonal cycle of sea surface temperature (SST) has emerged from the background noise of natural variability. Geographical patterns of changes in SST seasonal cycle amplitude (SST_{AC}) reveal two distinctive features: an increase at mid-latitudes in the Northern Hemisphere related to mixed-layer depth changes, and a robust dipole pattern between 40˚S and 55˚S in the Southern Hemisphere which is mainly driven by surface wind changes. The model-predicted pattern of SST_{AC} change is identifiable with high statistical confidence in four observed SST products and in 51 individual model realizations of historical climate evolution. Simulations with individual forcing reveal that greenhouse gas increases are the primary driver of changes in SST_{AC}, with smaller but distinct contributions from anthropogenic aerosol and ozone forcing. The robust human influence identified here on the seasonality of SST is likely to have wide-ranging impacts on marine ecosystems.

Earth's climate is simultaneously influenced by anthropogenic and natural external forcings, as well as by natural internal climate variability operating on a wide range of different space and time scales. Detection and attribution (D&A) analysis seeks to disentangle these human and natural influences[1]. Pattern-based "fingerprint" methods are a key component of D&A studies. Such methods have successfully identified human fingerprints in long-term annual-mean changes in surface and atmospheric temperature[2–7], different aspects of the hydrological cycle[8–12], atmospheric circulation[13,14], and ocean heat content[15,16].

The annual cycle is one of the most fundamental aspects of our climate and accounts for greater than 90% of seasonal temperature variability over most of the globe[17]. It influences human health, water supplies, agriculture, energy demand, and ecosystems. Gaining insight into how anthropogenic forcing has impacted seasonality is of scientific, economic, and societal importance.

Although annual cycle changes have attracted recent scientific attention in D&A studies[17–22], such investigations have not been performed with ocean variables.

We focus here on changes in the amplitude of the annual cycle of sea surface temperature (SST_{AC}), which plays an important role in air-sea interaction, global rainfall patterns, and the distributions of marine ecosystems[23–25]. In the tropical Pacific, model projections show an intensification of SST_{AC} in the twenty-first century compared to the twentieth century, which has been attributed to changes in meridional SST gradients[26] and atmospheric circulation[27]. In the mid-latitudes, SST_{AC} is projected to increase in both hemispheres[24,28,29]. These projections of SST_{AC} intensification in mid-latitudes are consistent with the observed amplitude increase in the surface air temperature (SAT) and tropospheric temperature (TT) annual cycles[17,22] during recent decades. Since SST, SAT, and TT are independently measured, the emergence of an externally forced signal in SST_{AC} would provide additional support for the identification of anthropogenic fingerprints in SAT and TT annual cycles.

Several previous model investigations demonstrate that the mid-latitude amplification of SST_{AC} is primarily linked to changes in mixed-layer depth (*MLD*)[24,28–31]. In summer, decreasing *MLD* leads to trapping of the net surface heat flux into the ocean in a thinner layer, thereby yielding a larger summertime SST increase[28]. This shoaling of the mixed layer results from enhanced upper ocean stratification driven by ocean warming[32,33]. In simulations with estimated future greenhouse-gas emissions, the annual-mean mixed layer shoaling and the mid-latitude SST_{AC} increase are projected to intensify[29] as the effective heat capacity of the thinner mixed layer decreases.

It is still unclear if an anthropogenic fingerprint can be formally detected in the changing amplitude of the observed SST annual cycle, and whether this fingerprint can be robustly attributed to human influence. We address this question here with four different observed SST datasets and over 50 individual model realizations of historical climate change. A novel aspect of our fingerprint study is its use of idealized simulations and heat budget analysis to elucidate the physical mechanisms that dictate key features of the common model and observed patterns of SST_{AC} change.

Trends in SST_{AC}

In all four of the observed SST products we examined, SST_{AC} trends over our primary analysis period (1950-2014) increase in most ocean regions and have a similar spatial pattern (Fig. 1a-d). Some features of the observed pattern are also evident in model simulations of historical climate change (HIST; Fig. 1e). The changes common to the models and observations are dominated by zonal-mean amplitude increases between 30˚ and 60˚ in both hemispheres (Fig. 1f), poleward of the maxima in the SST_{AC} climatology (Extended Data Fig. 1). Another notable regional-scale feature of the SST_{AC} trends in the HIST multi-model mean (MMM) and the HadISST and PCMDI observed datasets is the decrease in annual cycle amplitude in the vicinity of the Antarctic Circumpolar Current (Fig. 1e) south of 50˚S.

Although models can reproduce the positive observed SST_{AC} trends at NH mid-latitudes, the observed trends are smaller than in the simulations (Fig. 1f). One possible interpretation of this result is that the observed regional signals may be partly suppressed by the specific phasing of internal variability in the North Atlantic, as is the case with observed annual-mean warming in the tropical Pacific[34,35]. Differences between SST_{AC} trends in observed data and the HIST MMM are also prominent in the tropics, such as the pronounced maximum in the western equatorial Pacific that appears only in observations.

These model-observed differences may be partly due to the fact that the MMM is an average over individual realizations of historical climate change (in a single model) and an average over models. Averaging damps the noise of natural internally generated variability, which is uncorrelated across model realizations (except by chance). The MMM, therefore, should more clearly reveal the response to external forcing[34,35]. In contrast, there is only one realization of the observed record, which contains both internal variability and the forced signal in SST_{AC}. We therefore expect observed SST_{AC} changes to be noisier than in the MMM, particularly in regions where multidecadal variability affects tropical and subtropical temperature trends[34,35].

Could the above-mentioned model-observed differences in SST_{AC} trends be related to model biases in climatological-mean SST_{AC} patterns? We find that the model-average correlation between the patterns of model biases in the climatology of SST_{AC} and the model biases in the patterns of trends of SST_{AC} is low (R = 0.06). We infer from this that model biases in climatology do not appear to be a dominant factor in explaining the differences between the observed and simulated SST_{AC} trend patterns in Fig. 1a-d and Fig. 1e. This does not, however, rule out a possibility we discuss later – that overestimated climate sensitivity may contribute to model-observed differences in SST_{AC} trends[36].

The MMM and observations show closer agreement in global-scale features of zonally averaged SST_{AC} trends (Fig. 1f), with a common pattern of larger increase in the amplitude of SST_{AC} in the extratropics relative to the tropics. This pattern occurs in both hemispheres, but the mid-latitude increases in SST_{AC} are larger and broader in the Northern Hemisphere (NH) than in Southern Hemisphere (SH). This hemispheric asymmetry is consistent with results from previous studies of changes in the amplitude of the annual cycle of mid-tropospheric temperature[17,22]. As noted above, the simulated decrease of SST_{AC} trends in the Southern Ocean (Fig. 1f) is common to HadISST and PCMDI. Although the other two observational estimates do not show negative SST_{AC} trends between 50˚-60˚S, they have trend magnitudes within this latitude band that are smaller than the positive trends between 35˚-45˚S, and thus are consistent with the MMM results in a relative sense.

Fingerprint Analysis and Detection Time

We use a standard pattern-based method to determine whether the model-predicted externally-forced fingerprint of SST_{AC} changes is statistically identifiable in observations[37]. The fingerprint we search for is the leading empirical orthogonal function (EOF) of the MMM SST_{AC} anomalies (Methods). The fingerprint is calculated from the HIST simulations over the period 1950 to 2014 (Fig. 2a). Our analysis assumes that the spatial structure of the fingerprint pattern does not change markedly over time[17,38]. We tested and confirmed this assumption by calculating the HIST fingerprint for four different analysis periods (1950-2014, 1960-2014, 1970-2014, and 1980-2014; see Extended Data Fig. 2).

We compare the time-invariant SST_{AC} fingerprint pattern calculated from the HIST MMM with the time-evolving SST_{AC} patterns from observed datasets and long model control runs, respectively. These comparisons yield time series of similarity between the fingerprint and observed SST_{AC} patterns and between the fingerprint and patterns of natural internal variability in SST_{AC}. By varying the trend length L over a range of timescales (from 10 to 65 years), we can determine whether (and when) the similarity between the observations and the HIST fingerprint shows a statistically significant signal – i.e., an increase in pattern similarity over time that is unlikely to be due to natural internal variability alone.

Figure 2b shows the timescale-dependent S/N ratio calculated from the trends of these signal and noise time series. We stipulate that fingerprint detection occurs at trend length *L* if the S/N ratio exceeds a 5% significance threshold and remains above this threshold for all trend lengths larger than *L*. The model HIST fingerprint is identifiable with high statistical confidence (i.e., at the 5% significance level or better) in all four observational SST datasets after ca. 2000. At the end of the 65-year record, S/N ratios in the observations vary between 2.8 and 3.5. This indicates that smaller-scale differences between the four observational data sets – such as the previously-noted SST_{AC} trend differences at high latitudes in the SH – have relatively small impact on detection of the global-scale fingerprint in observations.

We also show the S/N ratios obtained when the HIST MMM fingerprint is searched for in individual realizations of HIST simulations (gray curves in Fig. 2b). In all 51 realizations, S/N exceeds the 5% threshold before the end of the simulation period in 2014. As in the case of the observations, SST_{AC} changes in individual HIST runs exhibit time-increasing similarity with the fingerprint (gray curves in Fig. 2b), pointing towards the robustness of the model-predicted forced SST_{AC} response.

The S/N ratios calculated with observed data generally lie within but close to the lower end of the model-generated S/N ratio distribution. There are multiple (not mutually exclusive) possible explanations for this result. These explanations include errors in the model external forcings[39], errors in the simulated SST_{AC} responses to the applied forcings, residual systematic errors in the observations, and model-versus-observed mismatches in the random phasing of internal variability (e.g., the El Niño–Southern Oscillation, Interdecadal Pacific Oscillation, and Pacific Decadal Oscillation). The latter explanation contributes to the more muted observed annual-mean tropospheric warming over the satellite era[34].

It is still unclear, however, what influence such mismatches in simulated and observed variability phasing have on changes in the seasonal cycle of SST. Here, we note that individual ensemble members generated with the same model and external forcings can have appreciable differences in their S/N behavior (see Extended Data Fig. 3). This suggests that as in the case annual-mean tropospheric temperature changes[34], model-observed differences in the phasing of internal variability may have marked influence on SST_{AC}, and hence on the overestimated "model only" S/N ratios in Fig. 2b. The non-negligible correlation between climate sensitivity[40] and the "model only" S/N ratios over the full 65-year analysis period (R = 0.55) provides evidence that overestimated model climate sensitivity[36] could also contribute to overestimated "model only" S/N ratios (see Extended Data Fig. 4).

Figure 2c provides information on the monitoring period required to identify the model-predicted HIST SST_{AC} fingerprint (the trend length L). Values of L at which detection occurs are shown as a function of the choice of the analyzed period. We consider four different periods; each ends in 2014, but has a different start date (1950, 1960, 1970, and 1980). There are two principal findings from this analysis. First, irrespective of the assumed start date of monitoring, the model-predicted HIST SST_{AC} fingerprint pattern in Fig. 2a is robustly identifiable in all four observed SST data sets and in all 51 model realizations of historical climate change. Except in the case of S/N results obtained with COBE data, the observed values of L are always contained within the spread of the model results.

Second, a common feature of both the simulated and observed results is that L decreases systematically with later start dates. For the MMM SST_{AC} changes, L is approximately 48 years and 18 years for start dates in 1950 and 1980 (respectively). This systematic decrease is likely due to larger net positive anthropogenic forcing over the 1980-2014 period than over periods with earlier start dates that sample appreciable negative forcing by anthropogenic aerosols. As will be shown in the next section, GHG forcing is the dominant influence on simulated SST_{AC} changes, so changes over time in the relative importance of GHG and anthropogenic aerosol forcing must contribute to the differences in L in the four analysis periods in Fig. 2c. Note that for fingerprint detection in the four different observed SST data sets, the spread in L values decreases as a function of increasing start date. This decrease in spread is partly due to improvements over time in the quality and spatial coverage of SST measurements and overlap between datasets.

Contributions from individual external forcings

We use single-forcing simulations to isolate and quantify the individual contributions of changes in well-mixed greenhouse gases (GHG), anthropogenic aerosols (AER), stratospheric ozone depletion (O3), and volcanic eruptions and solar variability (NAT) (Methods). We apply two different methods to understand the effects of single forcings: (1) To estimate the contributions of individual external forcings to the time-evolving S/N ratios obtained with the HIST MMM fingerprint, the GHG, AER, O3, and NAT single-forcing simulations are all regressed onto the same HIST fingerprint used in the previous section; (2) To determine whether the model-predicted fingerprint associated with an individual forcing is statistically identifiable, SST_{AC} changes from observations and HIST runs are projected onto each of the four fingerprints estimated from the GHG, AER, O3, and NAT single-forcing experiments. In the main text, we focus on Method 1 results. The results based on Method 2 are discussed in Methods.

The Method 1 S/N results indicate that GHG forcing is the dominant contributory factor to the identification of the HIST SST_{AC} fingerprint, which is detectable in the GHG MMM before 1990 (and before the end of the analysis period in 2014 in 48 out of 51 individual GHG realizations; see Extended Data Fig. 5). The S/N ratios for the "GHG only" case increase nearly linearly with increases in timescale L and the magnitude of the GHG forcing. In contrast, S/N results for AER show markedly nonlinear behavior as L increases. This is due to non-monotonic changes in emissions of anthropogenic sulfate aerosols, with large emissions after World War II followed by a reduction in emissions from North America and Europe after the 1980s[41–43]. The HIST SST_{AC} fingerprint is not detectable in the MMM of AER, O3, or NAT.

Our analysis of the impact of individual anthropogenic factors assumes additivity of the forced responses in GHG, AER, O3, and NAT[44,45]. To test the validity of this additivity assumption, we compare the HIST S/N results in Fig. 3a with S/N results obtained for ALL, the linear combination of the individual S/N ratios obtained for the GHG, AER, O3 and NAT experiments. Additivity is a reasonable assumption for analysis periods longer than 40 years. For periods less than 40 years, differences between the HIST and ALL S/N results are likely related to the combined effects of larger noise on shorter timescales, the smaller ensemble size for O3, and nonlinear aspects of the forced SST_{AC} responses[46–48].

Figure 3b provides detection times for the HIST SST_{AC} fingerprint in HIST, GHG, and three linear combinations of individual SST_{AC} responses: GHG+AER, GHG+O3, and GHG+AER+O3. The primary influence on detection time is GHG, with AER acting to delay fingerprint detection: "GHG only" yields systematically earlier detection times than any set of SST_{AC} changes that includes AER (HIST, GHG+AER, or GHG+AER+O3). Including O3 also advances detection time, with the earliest median detection time of the HIST SST_{AC} fingerprint (in 1985) in the GHG+O3 linear combination. The spread in detection times obtained with linear combinations is larger than the spread in detection time inferred from HIST. This is likely due to amplification of noise in the linear combination of individual responses.

Physical drivers of SST changes

We seek to understand the physical drivers of the SST_{AC} changes described in the previous sections. In the observations, warming of zonal-mean SST over 1950 to 2014 occurs in nearly all months and latitudes (Extended Data Fig. 6). For the mid-latitudes it is more pronounced in the summer hemisphere. In the SH at ca. 40˚S, both the observations and HIST display warming relative to annual-mean trends in austral summer and cooling relative to annual-mean trends in austral winter (Figs. 4a,b). In HIST, this feature is primarily driven by GHG forcing (Figs. 4c). Relative to observations, CMIP6 models yield larger NH temperature rises in both summer and winter. As noted above, there are multiple possible interpretations of this result.

Another prominent aspect of HIST and GHG is a dipole pattern characterized by anticorrelation between the seasonal temperature changes at roughly 40˚S and 55˚S. GHG and O3 forcing both contribute to this feature (Figs. 4c,d). As noted above, this dipole is evident in two of the four observed datasets (HadISST and PCMDI; Extended Data Fig. 7). These observational differences likely arise because satellite data were included in HadISST and PCMDI but not in ERSST and COBE. In consequence, the Southern Ocean is better represented in the first two datasets, especially in the vicinity of sea-ice.

Buoyancy flux and wind stress changes are two major surface forcings affecting the Southern Ocean climate[49–51]. We explore the respective effects of buoyancy (dominated by heat flux change) and wind (momentum) forcing on SST_{AC} changes using the Flux-Anomaly-Forced Model Intercomparison (FAFMIP) experiments (Fig. 5). In the FAF-stress experiment, in which CO_2-induced perturbations to the ocean are imposed in wind stress only, the SH mid-latitudes show a robust meridional dipole pattern in zonal-mean SST_{AC} change (Fig. 5b). In the FAF-heat experiment, CO_2-driven perturbations to heat fluxes amplify SST_{AC} in both hemispheres, but the magnitude of the change is markedly larger in the NH (Fig. 5c), where the wind stress effect is limited. The FAFMIP results imply that wind forcing caused by CO_2 increases is the main driver of the above-described SST_{AC} dipole pattern between 40˚S and 55˚S found in HIST, GHG, and two of the observed SST datasets. In contrast, changes in NH mid-latitude SST_{AC} arise from increased surface heat flux linked to atmospheric warming.

In addition to the influence of these surface wind stress and heat flux forcings, the SST_{AC} fingerprint can also be influenced by ocean adjustments arising from *MLD* changes. We investigated the role of *MLD* changes with a simplified mixed-layer heat budget analysis of the HIST runs. Our heat budget model also considers the effects of net surface heat flux (*Qnet*) and shortwave radiation flux out of the mixed layer base into the intermediate ocean (*Qb*) (Methods; Eq. 3). The patterns of the $dSST/dt_{AC}$ change can be reproduced by this simple model (Figs. 6a,b), and are consistent with the SST_{AC} fingerprint (Fig. 2a). The shoaling of *MLD* with fixed *Qnet-Qb* is the key factor here (Fig. 6c). In winter, this shoaling effect generates SST cooling by enhancing temperature response to winter heat loss. In summer, shoaling yields SST warming. It is noteworthy that the *Qnet* effect here differs from the analysis of FAF-heat, as the latter also incorporates *MLD* effects arising from the accumulated ocean heat.

Formatted: Font: Italic

Because of this seasonally dependent effect of the *MLD* shoaling, the SST_{AC} would be amplified even with constant *MLD* shoaling throughout the year. This is why both hemispheres show positive annual cycle changes in the 30˚-50˚ latitude band. Between 50˚S-60˚S, the *MLD* deepens in austral summer, which appears to overwhelm the shoaling of *MLD* in austral winter, thus decreasing SST_{AC} in this band. The fixed *MLD* case results in a weak but reduced SST_{AC} in most regions (Fig. 6d), which implies that the warming induced by the *Qnet-Qb* change is slightly larger in winter than in summer.

We performed two further sensitivity experiments: (1) constant monthly *MLD* shoaling, in which the summer value is applied for all 12 months at each location; and (2) shoaling *MLD* by 5% in every month and location relative to the background monthly value. Our results suggest that the absolute change and relative change of *MLD* give rise to similar patterns (Extended Data Fig. 8). The major difference is in the 50˚S-60˚S band, apparently due to the opposite directions of *MLD* change between austral winter and summer (Extended Data Fig. 9a,b). For all other latitudes, the shoaling of the mixed layer is consistent with season.

The westerly wind stress in the 50˚S-60˚S region increases in austral summer (Extended Data Fig. 9c,d). This can deepen *MLD* by increased local turbulent mixing as well as by the increased equatorward advection of colder water. The negative wind stress changes between 30˚S-50˚S have the opposite effect. The contrasting surface wind responses in the 30˚S-50˚S and 50˚S-60˚S bands reflect the poleward shifting of zonal winds over the Southern Ocean caused by GHG and O3 forcing (Fig. 4c,d). This shift is consistent with the FAF-stress response to CO_2-driven wind stress changes.

Conclusions

Most previous studies of the annual cycle of SST (SST_{AC}) focused primarily on projected 21st century changes[29,31]. Here, we examine whether there is a detectable "fingerprint" pattern of human-induced SST_{AC} change over 1950-2014. We provide the first scientific evidence that a human-caused SST_{AC} signal has already emerged from the background noise of natural variability. Geographical patterns of SST_{AC} changes show increased SST_{AC} at mid-latitudes in the NH and a distinctive meridional dipole structure at SH mid-latitudes. These large-scale zonal features are common to observations and model simulations with anthropogenic forcing, and are dissimilar to the smaller-scale structure of natural internal variability. This helps to explain why the model-estimated SST_{AC} fingerprint in response to combined anthropogenic and natural external forcing is identifiable by the end of the 20th century in all four observed SST datasets analyzed here. The fingerprint is also robustly identifiable in all 51 model realizations of historical climate change.

Single forcing experiments indicate that increases in well-mixed GHGs is the dominant factor in the identification of externally forced changes in SST_{AC}. Anthropogenic aerosol emissions are likely to have delayed the detection of this fingerprint by ca. 7-8 years on average. External forcing from stratospheric ozone depletion partially contributed to the development of the SST_{AC} dipole structure at SH mid-latitudes, while natural external forcing by volcanoes and solar irradiance changes had relatively little effect on the detection of a human fingerprint in SST_{AC}.

Deleted: are the main driver of the large-scale spatial structure of the model-predicted SST_{AC} fingerprint

Model simulations and a heat budget analysis reveal that the leading physical drivers of these large-scale SST_{AC} changes are different in the two hemispheres. In the SH, the impacts of changes in atmospheric circulation and surface wind stress on the *MLD* are the key determinant of the dipole-like SST_{AC} response in the Southern Ocean. In the mid-latitudes of both hemispheres, human-induced warming yields increased stratification of the upper ocean, which in turn causes shoaling of the *MLD* during all seasons. Year-round *MLD* shoaling decreases the thermal inertia, thereby amplifying the mid-latitude SST_{AC}.

Human-driven amplification of the mid-latitude seasonal cycle of SST has important implications for future changes in the behavior of marine ecosystems. The SST changes found here have the potential to influence both the productivity and distribution of marine species which constitute key food resources for human societies. Our finding of robust changes in the seasonality of SST should motivate more detailed exploration of the anthropogenically forced seasonal changes in a wide range of different ocean properties.

Methods

CMIP6 experiments and models.

This study uses output from climate model simulations performed under Phase 6 of the Coupled Model Intercomparison Project (CMIP6)[52]. We focus on 10 CMIP6 models that performed all of the following 4 experiments: historical all-forcing simulations (HIST) and single-forcing simulations performed with anthropogenic aerosols (AER), greenhouse gases (GHG), and purely natural changes in solar irradiance and volcanic aerosols (NAT)[53]. Each of the 10 models has multiple ensemble members. Each ensemble member of a given model is driven by the same external forcing, but has a different manifestation of natural internal climate variability superimposed on the underlying forced response. The number of ensemble members available for each model and each experiment is listed in Extended Data Table 1. For each experiment, there are 51 realizations in total. The multi-model mean (MMM) is the average of the ensemble means of these 10 models. The preindustrial control (piControl) simulations from the same 10 models are used for the purpose of estimating the noise from internal variability (see below).

We also analyze results from an experiment with forcing by stratospheric ozone changes only (O3). Only 4 of the 10 models that performed HIST, GHG, AER, and NAT simulations provided results for the O3 simulation (see Extended Data Table 1).

The HIST, GHG, AER, O3, and NAT experiments cover the period from 1850 to 2014. We focus on the 1950-2014 period for comparing simulations with observations of changes in the amplitude of the annual cycle of SST (SST_{AC}). This choice of period was dictated by improvement in the spatial coverage and quality of observed SST data after World War II, as well as by large post-1950 changes in well-mixed GHGs, anthropogenic aerosols, and stratospheric ozone. All model output was interpolated to a common, regular $1^\circ \times 1^\circ$ grid.

Observations.

We rely on four primary SST gridded products. These are the Hadley Center Sea Ice and SST dataset version 1 (HadISST)[54], the NOAA Extended Reconstructed SST dataset version 5 (ERSST)[55], the Centennial In Situ Observation-Based Estimates of the Variability of SST and Marine Meteorological Variables, version 2 (COBE)[56], and the Program for Climate Model Diagnosis and Intercomparison SST dataset (PCMDI)[57]. ERSST and COBE are based on in situ measurements, and HadISST and PCMDI combine in situ and satellite estimates of SST. Different averaging and gap-filling approaches are employed to infill data-sparse regions and time periods in these gridded products. HadISST and PCMDI datasets are not entirely independent: the PCMDI dataset is HadISST1 through 1981, and uses the NOAA Optimum Interpolation SST data (OI.v2)[58] thereafter.

In addition to these observational SST products, we also used the monthly surface zonal wind from the latest-generation reanalysis of the European Centre for Medium-Range Weather Forecasts (ERA5)[59]. For the observed mixed layer depth (*MLD*), we first employed the gridded monthly temperature and salinity data from the IAP product[60] to calculate the potential density. *MLD* was then defined as the depth at which the ocean potential density exceeds the sea surface density at a criterion of $\delta\rho = 0.125$ kg/m³, following the definition for *MLD* output (referred to as 'mlotst') from the CMIP6 models. There are likely to be substantial uncertainties in the IAP product arising from sparse measurements of the subsurface temperature and salinity fields in the Southern Ocean (particularly in the pre-Argo era of the IAP records).

In addition, we have used the information from ref 61 to examine whether biases in ship SST data could be an important factor in our D&A analysis. We find it is unlikely that ship SST data biases could alter any of our findings regarding the identification of an SST_{AC} fingerprint in observations (not shown).

FAFMIP experiments.

To isolate the individual effects of changes in wind stress and surface heat flux on SST_{AC} trends, we rely on output from the Flux-Anomaly-Forced Model Intercomparison (FAFMIP) experiments. Results are from 5 models: ACCESS-CM2, CanESM5, HadGEM3-GC31-LL, MIROC6, MRI-ESM2-0. The FAFMIP experiments, branched from each model's piControl run, prescribe a set of surface flux perturbations for the ocean. These perturbations are obtained from the ensemble-mean changes simulated at the time of doubled CO_2 by CMIP5 AOGCMs run under the 1pctCO2 scenario (in which atmospheric CO_2 levels increase by 1% each year). We examine three different FAFMIP experiments: FAF-all, in which perturbations of surface wind stress, surface freshwater flux, and surface heat flux are simultaneously imposed; FAF-stress, with imposed perturbations of surface wind stress only; and FAF-heat, with imposed perturbations of net surface heat flux only[62].

All FAFMIP experiments considered here were run for 70 years. We show the anomalies of the 31–70-yr average relative to the climatology calculated from the full length of each model's piControl.

Calculation of annual cycle amplitudes.

For each model simulation and observation product, and at each grid-point *x* and year *t*, we performed a Fourier analysis on the 12 monthly-mean values of SST. The amplitude of the first harmonic is taken as the annual cycle amplitude (SST_{AC}; see Extended Data Fig. 1d). Consistent with previous work[17], the first harmonic explains > 95% of the total seasonal variance at almost all locations between 60˚N and 60˚S (except at regions close to the equator). As an additional sensitivity study, we confirmed that our fingerprint results are insensitive to the definition of SST_{AC}. The S/N ratios and detection times obtained here with the first harmonic are very similar to those found when we define SST_{AC} as the seasonal maximum SST minus the seasonal minimum SST at each grid-point and in each year.

Pattern-based fingerprint analysis.

***a.* Definition of the fingerprint**

Detection methods generally require an estimate of the true but unknown climate-change signal, typically designated as the fingerprint $F(x)$, in response to an individual forcing or set of forcings[63]. As in previous work, we assume $F(x)$ to be the first EOF of the MMM change in SST_{AC} in the HIST simulations[17].

Let $S(i, j, x, t)$ represent SST_{AC} at grid-point x and year t from the i^{th} realization of the j^{th} model's HIST simulation, where:

$i = 1, \ldots N_r(j)$ (the number of realizations for the j^{th} model)
$j = 1, \ldots N_m$ (the number of models used in fingerprint estimation)
$x = 1, \ldots N_x$ (the total number of grid-points after regridding to a regular 1˚× 1˚grid)
$t = 1, \ldots N_t$ (the time in years)

Here, N_r varies across models (Extended Data Table 1). For HIST, $N_m = 10$ models. Prior to the fingerprint analysis, all model and observed SST_{AC} fields were interpolated to a common 1˚× 1˚ latitude/longitude grid. The evolution of multi-model mean SST_{AC} was calculated by first averaging over an individual model's realizations (where multiple realizations were available), and then averaging across the number of models available for each experiment. MMM anomalies were then defined at each grid-point x and year t with respect to the local MMM climatological annual cycle amplitude. The fingerprint is the first EOF of the changes over time in the MMM SST_{AC} anomalies from the HIST experiment – i.e., the temporal changes in the annual cycle of SST. To minimize the impact of sea-ice on SST_{AC}, the domain was restricted to 60˚N-60˚S and to regions where the winter sea-ice concentration is smaller than 10%. The anomalies are weighted by the square root of the cosine of the grid node's latitude[64] before calculating the EOF. Most of the discussion focuses on model fingerprints estimated over 1950 to 2014. We also calculated fingerprints for three additional analysis periods (1960-2014, 1970-2014, and 1980-2014). As noted in the main text, the spatial structure of the fingerprint patterns does not change markedly over these periods (Extended Data Fig. 2).

***b.* Fingerprint detection**

We seek to determine whether the pattern similarity between the time-varying observations and $F(x)$ shows a statistically significant increase over time. To address this question, we require control run estimates of internally generated variability ("noise"), in which we know *a priori* that there is no expression of the fingerprint, except by chance.

This intrinsic noise is estimated using preindustrial control runs (piControl) from the same 10 models employed for calculating the HIST fingerprint. These control simulations can be affected by residual long-term drift. To reduce the effects of such drift on estimates of the internal variability of SST_{AC}, we fit a cubic polynomial to the full length of each model's control run and then removed the fitted polynomial[65,66]. Fitting and drift removal is performed at each model grid-point. Because the individual model control runs are of unequal length, our noise estimates rely on the last 400 years of each model's piControl run. This yields a total of 4,000 years of concatenated control run data, and avoids introducing any bias associated with differing control run lengths.

Observed SST_{AC} estimates are expressed as anomalies relative to climatological means over the 1950-2014 analysis period (or over the alternate analysis periods in Fig. 2c). The observed temperature data are projected onto *F(x)*, the time-invariant fingerprint:

$$Z_o(t) = \sum_{x=1}^{N_x} O(x,t)\, F(x) \tag{1}$$

where *O(x, t)* are the observed SST_{AC} anomalies. This projection is equivalent to a spatially uncentered covariance between the patterns *O(x, t)* and *F(x)* at year *t*. The signal time series $Z_o(t)$ provides information on the fingerprint strength in the observations. If observed patterns of temperature change are becoming increasingly similar to *F(x)*, $Z_o(t)$ should increase over time.

To assess whether this increase is statistically significant, we compare trends in $Z_o(t)$ with a null distribution for which we know *a priori* that there is no expression of the fingerprint, except by chance. Here, we derive this null distribution using *C(x, t)*, the 4,000-year concatenated noise data set, generated from the piControl runs as described above. The noise time series $N_c(t)$ is the projection of *C(x, t)* onto the fingerprint:

$$N_c(t) = \sum_{x=1}^{N_x} C(x,t)\, F(x) \tag{2}$$

where the length of $N_c(t)$ is 4,000 years (see above).

We estimate signal-to-noise (S/N) ratios by fitting least-squares linear trends of increasing length *L* years to $Z_o(t)$, and then comparing these trends with the standard deviation of the distribution of maximally overlapping *L*-length trends in $N_c(t)$[17,37]. Signal detection is stipulated to occur when the trend in $Z_o(t)$ exceeds and remains above the stipulated significance level (which is 5% in our study)[22]. The test is one-tailed, and we assume a Gaussian distribution of trends in $N_c(t)$. The start date for fitting linear trends to $Z_o(t)$ is 1950 for our baseline analysis, and is 1960, 1970, and 1980 in the alternate analysis periods shown in Fig 2c. We use a minimum trend length of ten years, so the first S/N ratio (and the earliest possible detection time in the baseline period) is for 10-year trends ending in 1959.

We also show S/N results that are based solely on model simulation output. In our "model only" results, $N_c(t)$ is calculated as in Eq. 2, but the observational estimates in Eq. 1 are replaced by $S(i, j, x, t)$, the annual cycle amplitude information from each of the 51 HIST simulations (see the gray curves in Figs. 1f and 2b).

c. HIST fingerprint vs. single-forcing fingerprints

As noted in the main text, we employ two methods to study the contributions of individual external forcings (GHG, AER, NAT, and O3) to the simulated SST_{AC} changes. In Method 1, SST_{AC} anomalies from individual realizations of the four single-forcing simulations are projected onto the common fingerprint calculated from the HIST MMM. As in the case of HIST, the MMMs of SST_{AC} from these four single-forcing experiments were also projected onto $F(x)$.

In Method 2, we project SST_{AC} changes from the HIST MMM and from individual HIST realizations onto each of the four fingerprints estimated from the GHG, AER, O3, and NAT multi-model average SST_{AC} changes. This yields information on the strength of each individual fingerprint in the historical all-forcing simulations, and on how the strength of the GHG, AER, O3, and NAT fingerprints evolves with increasingly longer analysis periods.

We use EOF1 for the Method 2 GHG fingerprint and EOF2 for the Method 2 fingerprints from AER, O3, and NAT (see Extended Data Fig. 10). This choice was made because in the GHG simulation, EOFs 1 and 2 are clearly separated in terms of explained variance (EV), with the EV associated with GHG EOF1 a factor of 3 larger than the EV of GHG EOF2. The latter pattern largely reflects tropical internal variability associated with the El Niño-Southern Oscillation (ENSO). In contrast, EOFs 1 and 2 are less well separated in terms of EV in the AER, O3, and NAT simulations – their EOF1 is very similar to EOF2 from the GHG simulation, while the EOF2 patterns of AER, O3, and NAT appear to be dominated by extratropical forced responses. Moreover, the second principal component from AER, O3, and NAT show some remarkable signals in the temporal evolution (Extended Data Fig. 10j).

SST_{AC} from observations and HIST runs are projected onto these four single-forcing fingerprints (Extended Data Fig. 11). For the projections onto the GHG fingerprint, all 51 model HIST realizations and three of the four observational datasets eventually exceed the 5% significance threshold. S/N levels are systematically lower for the AER, NAT, and O3 fingerprints, which are therefore not as clearly identifiable in the HIST realizations or observations as the GHG fingerprint. This provides support for a key finding from our Method 1 analysis: forcing by well-mixed GHGs is the dominant factor in the identification of externally forced changes in SST_{AC}.

We note that in Method 2, the NAT fingerprint is identifiable at the 5% level in 88% of the HIST realizations and in two of the four observed SST_{AC} datasets (Extended Data Fig. 11d). While there are small changes over time in the solar and volcanic forcing over 1950 to 2014[39], the behavior of the first principal component of the NAT SST_{AC} changes (Extended Data Fig. 10i) suggests that NAT forcing is unlikely to produce a significant multi-decadal trend in SST_{AC}. Instead, the identification of the NAT fingerprint in the HIST SST_{AC} data appears to be due to the spatial similarity between certain large-scale features of the GHG and NAT fingerprints (compare Extended Data Figs. 10a,f). Thus in Method 2 (which we do not focus on in our fingerprint analysis) the statistical problem of degeneracy[67] of the normalized GHG and NAT fingerprints hampers reliable assessment of the relative contributions of GHG and NAT forcing to the simulated SST_{AC} changes. In Method 1, however, the larger amplitude of the SST_{AC} response to GHG forcing (relative to NAT forcing) is preserved – which is why the HIST fingerprint can be identified in the individual GHG realizations, but not in the individual NAT realizations.

The uncentered pattern correlation between GHG EOF1 and NAT EOF2 is higher than the pattern correlations between GHG EOF1 and the EOF2 patterns of other single-forcing experiments (Extended Data Table 2). This similarity may arise from major tropical volcanic eruptions in the 1950-2014 analysis period (Agung, El Chichón, and Pinatubo) and the associated shifts of the intertropical convergence zone[68], which in turn could affect the latitudinal location of regions of mid-latitude increases in SST_{AC}.

Simplified Mixed-Layer Heat Budget Analysis.

Our mixed-layer heat budget model is a simplified version of the traditional mixed-layer heat budget model that takes into account only the dominant heat fluxes and mixed-layer depth affecting the temperature of the oceanic mixed layer:

$$\frac{dT}{dt} \sim \frac{Qnet - Qb}{rho * Cp * MLD} \tag{3}$$

The left-hand side is the ocean temperature tendency, and the right-hand side is the estimate based on net surface heat flux (*Qnet*), shortwave radiation flux leaving the mixed-layer base (*Qb*), and mixed-layer depth (*MLD*). These terms are functions of month, latitude and longitude and are calculated from HIST runs. The terms *rho* and *Cp* are the density and specific heat of seawater, respectively.

For the changes in the annual cycle (*AC*) amplitude of *dT/dt*:

$$AC\left(\frac{dT_2}{dt}\right) - AC\left(\frac{dT_1}{dt}\right) \approx AC\left(\frac{Qnet_2 - Qb_2}{rho * Cp * MLD_2}\right) - AC\left(\frac{Qnet_1 - Qb_1}{rho * Cp * MLD_1}\right) \tag{4}$$

Deleted: not shown

where “1” represents the average of the period 1950-1979, and “2” represents the average of the period 1985-2014. The changes are the difference between these two 30-year periods. We also hold *Qnet* and *MLD* constant in Eq. 4 to isolate the effects due to *MLD* change and *Qnet* change:

$$\Delta MLD\ effect = AC\left(\frac{Qnet_1 - Qb_1}{rho * Cp * MLD_2}\right) - AC\left(\frac{Qnet_1 - Qb_1}{rho * Cp * MLD_1}\right) \tag{5}$$

$$\Delta Qnet\ effect = AC\left(\frac{Qnet_2 - Qb_2}{rho * Cp * MLD_1}\right) - AC\left(\frac{Qnet_1 - Qb_1}{rho * Cp * MLD_1}\right) \tag{6}$$

In Eqs. 5 and 6, the *Qb* and *Qnet* terms are for the same analysis period. Results are insensitive to whether *Qb* is chosen from period 1 or period 2.

We examine the effect of *MLD* change in terms of its absolute change (Eq. 7) and relative change (Eq. 8). As shown in Eq. 7, we assumed a summer *MLD* change to be added to all the months from the base period. In terms of relative change, we assumed *MLD* is assumed to shoal by 5% everywhere and in every month relative to the background value (Eq. 8).

$$\Delta MLD_{summer}\ effect = AC\left(\frac{Qnet_1 - Qb_1}{rho * Cp * (MLD_1 + (MLD_{2,summer} - MLD_{1,summer})}\right) - AC\left(\frac{Qnet_1 - Qb_1}{rho * Cp * MLD_1}\right) \tag{7}$$

$$\Delta MLD_{5\% shoaling}\ effect = AC\left(\frac{Qnet_1 - Qb_1}{rho * Cp * (MLD_1 * 0.95)}\right) - AC\left(\frac{Qnet_1 - Qb_1}{rho * Cp * MLD_1}\right) \tag{8}$$

Data availability

The CMIP6 historical, single-forcing, and FAFMIP simulation outputs are available on the Earth System Grid of the Program for Climate Model Diagnosis and Intercomparison (PCMDI): (https://esgf-node.llnl.gov/search/cmip6/). HadISST data are available at: https://www.metoffice.gov.uk/hadobs/hadisst. ERSST data are available at: https://www.ncei.noaa.gov/products/extended-reconstructed-sst. COBE data are available at: https://psl.noaa.gov/data/gridded/data.cobe2.html. PCMDI data are available at: https://doi.org/10.22033/ESGF/input4MIPs.16921. ERA5 data are available at: https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5. IAP data are available at: https://climatedataguide.ucar.edu/climate-data/ocean-temperature-analysis-and-heat-content-estimate-institute-atmospheric-physics.

Acknowledgments

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which led the design of CMIP6 and coordinated the work, and we also thank individual climate modeling groups (listed in Extended Data Table 1) for their efforts in performing all model simulations analyzed here.

Author contributions

J.-R.S. and B.D.S. conceived the study. J.-R.S. conducted the analysis and wrote the first draft. J.-R.S., B.D.S, Y.-O.K., and S.E.W. contributed to interpreting the results, writing, and editing the manuscript.

Competing Interests

The authors declare no competing interests.

References

1. Bindoff, N. L. *et al.* Detection and Attribution of Climate Change: from Global to Regional. in *Climate Change 2013 – The Physical Science Basis* (ed. Intergovernmental Panel on Climate Change) vol. 9781107057 867–952 (Cambridge University Press, 2014).
2. Hegerl, G. C. *et al.* Detecting Greenhouse-Gas-Induced Climate Change with an Optimal Fingerprint Method. *Journal of Climate* **9**, 2281–2306 (1996).
3. Santer, B. D. *et al.* A search for human influences on the thermal structure of the atmosphere. *Nature* **382**, 39–46 (1996).
4. Tett, S. F. B., Mitchell, J. F. B., Parker, D. E. & Allen, M. R. Human Influence on the Atmospheric Vertical Temperature Structure: Detection and Observations. *Science* **274**, 1170–1173 (1996).
5. Stott, P. A. *et al.* External Control of 20th Century Temperature by Natural and Anthropogenic Forcings. *Science* **290**, 2133–2137 (2000).
6. Santer, B. D. *et al.* Quantifying stochastic uncertainty in detection time of human-caused climate signals. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 19821–19827 (2019).
7. Santer, B. D. *et al.* Exceptional stratospheric contribution to human fingerprints on atmospheric temperature. *Proceedings of the National Academy of Sciences* **120**, 1–11 (2023).
8. Stott, P. A., Sutton, R. T. & Smith, D. M. Detection and attribution of Atlantic salinity changes. *Geophysical Research Letters* **35**, L21702 (2008).
9. Santer, B. D. *et al.* Incorporating model quality information in climate change detection and attribution studies. *Proceedings of the National Academy of Sciences* **106**, 14778–14783 (2009).
10. Terray, L. *et al.* Near-surface salinity as nature's rain gauge to detect human influence on the Tropical water cycle. *Journal of Climate* **25**, 958–977 (2012).
11. Pierce, D. W., Gleckler, P. J., Barnett, T. P., Santer, B. D. & Durack, P. J. The fingerprint of human-induced changes in the ocean's salinity and temperature fields. *Geophysical Research Letters* **39**, L21704 (2012).
12. Marvel, K. & Bonfils, C. Identifying external influences on global precipitation. *Proceedings of the National Academy of Sciences* **110**, 19301–19306 (2013).
13. Gillett, N. P., Fyfe, J. C. & Parker, D. E. Attribution of observed sea level pressure trends to greenhouse gas, aerosol, and ozone changes. *Geophysical Research Letters* **40**, 2302–2306 (2013).
14. Christidis, N. & Stott, P. A. Changes in the geopotential height at 500 hPa under the influence of external climatic forcings. (2015) doi:10.1002/2015GL066669.Received.
15. Barnett, T. P. *et al.* Penetration of Human-Induced Warming into the World's Oceans. *Science* **309**, 284–287 (2005).
16. Bilbao, R. A. F., Gregory, J. M., Bouttes, N., Palmer, M. D. & Stott, P. Attribution of ocean temperature change to anthropogenic and natural forcings using the temporal, vertical and geographical structure. *Climate Dynamics* **53**, 5389–5413 (2019).
17. Santer, B. D. *et al.* Human influence on the seasonal cycle of tropospheric temperature.

Science **361**, (2018).

18. Min, S.-K., Zhang, X., Zwiers, F. W. & Agnew, T. Human influence on Arctic sea ice detectable from early 1990s onwards. *Geophysical Research Letters* **35**, L21701 (2008).
19. Qian, C. & Zhang, X. Human Influences on Changes in the Temperature Seasonality in Mid- to High-Latitude Land Areas. *Journal of Climate* **28**, 5908–5921 (2015).
20. Marvel, K. *et al.* Observed and Projected Changes to the Precipitation Annual Cycle. *Journal of Climate* **30**, 4983–4995 (2017).
21. Duan, J. *et al.* Detection of human influences on temperature seasonality from the nineteenth century. *Nature Sustainability* **2**, 484–490 (2019).
22. Santer, B. D. *et al.* Robust anthropogenic signal identified in the seasonal cycle of tropospheric temperature. *Journal of Climate* 1–51 (2022) doi:10.1175/jcli-d-21-0766.1.
23. Dwyer, J. G., Biasutti, M. & Sobel, A. H. The effect of greenhouse gas-induced changes in SST on the annual cycle of zonal mean tropical precipitation. *Journal of Climate* **27**, 4544–4565 (2014).
24. Alexander, M. A. *et al.* Projected sea surface temperatures over the 21st century: Changes in the mean, variability and extremes for large marine ecosystem regions of Northern Oceans. *Elementa* **6**, (2018).
25. Kwiatkowski, L. *et al.* Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. *Biogeosciences* **17**, 3439–3470 (2020).
26. Timmermann, A., Jin, F. F. & Collins, M. Intensification of the annual cycle in the tropical Pacific due to greenhouse warming. *Geophysical Research Letters* **31**, 1–4 (2004).
27. Sobel, A. H. & Camargo, S. J. Projected future seasonal changes in tropical summer climate. *Journal of Climate* **24**, 473–487 (2011).
28. Chen, C. & Wang, G. Role of North Pacific mixed layer in the response of SST annual cycle to global warming. *Journal of Climate* **28**, 9451–9458 (2015).
29. Jo, A. R. *et al.* Future Amplification of Sea Surface Temperature Seasonality Due To Enhanced Ocean Stratification. *Geophysical Research Letters* **49**, 1–10 (2022).
30. Liu, F., Lu, J., Luo, Y., Huang, Y. & Song, F. On the oceanic origin for the enhanced seasonal cycle of SST in the midlatitudes under global warming. *Journal of Climate* **33**, 8401–8413 (2020).
31. Dwyer, J. G., Biasutti, M. & Sobel, A. H. Projected changes in the seasonal cycle of surface temperature. *Journal of Climate* **25**, 6359–6374 (2012).
32. Capotondi, A., Alexander, M. A., Bond, N. A., Curchitser, E. N. & Scott, J. D. Enhanced upper ocean stratification with climate change in the CMIP3 models. *Journal of Geophysical Research: Oceans* **117**, n/a-n/a (2012).
33. Li, G. *et al.* Increasing ocean stratification over the past half-century. *Nature Climate Change* **10**, 1116–1123 (2020).
34. Po-Chedley, S. *et al.* Natural variability contributes to model–satellite differences in tropical tropospheric warming. *Proceedings of the National Academy of Sciences of the United States of America* **118**, 1–7 (2021).
35. Po-Chedley, S. *et al.* Internal variability and forcing influence model–satellite differences in the rate of tropical tropospheric warming. *Proceedings of the National Academy of*

Sciences **119**, 1–8 (2022).
36. Zelinka, M. D. *et al.* Causes of Higher Climate Sensitivity in CMIP6 Models. *Geophysical Research Letters* **47**, 1–12 (2020).
37. Santer, B. D. *et al.* Identifying human influences on atmospheric temperature. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 26–33 (2013).
38. Solomon, S. *et al.* Emergence of healing in the Antarctic ozone layer. *Science* **353**, 269–274 (2016).
39. Fyfe, J. C., Kharin, V. V., Santer, B. D., Cole, J. N. S. & Gillett, N. P. Significant impact of forcing uncertainty in a large ensemble of climate model simulations. *Proceedings of the National Academy of Sciences of the United States of America* **118**, 1–6 (2021).
40. Meehl, G. A. *et al.* Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances* **6**, 1–10 (2020).
41. Deser, C. *et al.* Isolating the Evolving Contributions of Anthropogenic Aerosols and Greenhouse Gases: A New CESM1 Large Ensemble Community Resource. *Journal of Climate* **33**, 7835–7858 (2020).
42. Shi, J.-R., Kwon, Y.-O. & Wijffels, S. E. Two Distinct Modes of Climate Responses to the Anthropogenic Aerosol Forcing Changes. *Journal of Climate* **35**, 3445–3457 (2022).
43. Shi, J.-R., Kwon, Y. & Wijffels, S. E. Subsurface Ocean Temperature Responses to the Anthropogenic Aerosol Forcing in the North Pacific. *Geophysical Research Letters* **50**, 1–16 (2023).
44. Ribes, A. & Terray, L. Application of regularised optimal fingerprinting to attribution. Part II: Application to global near-surface temperature. *Climate Dynamics* **41**, 2837–2853 (2013).
45. Marvel, K. *et al.* Do responses to different anthropogenic forcings add linearly in climate models? *Environmental Research Letters* **10**, (2015).
46. Maher, N., McGregor, S., England, M. H. & Gupta, A. Sen. Effects of volcanism on tropical variability. *Geophysical Research Letters* **42**, 6024–6033 (2015).
47. O'Brien, J. P. & Deser, C. Quantifying and Understanding Forced Changes to Unforced Modes of Atmospheric Circulation Variability over the North Pacific in a Coupled Model Large Ensemble. *Journal of Climate* **36**, 19–37 (2023).
48. Deser, C., Simpson, I. R., Phillips, A. S. & McKinnon, K. A. How Well Do We Know ENSO's Climate Impacts over North America, and How Do We Evaluate Models Accordingly? *Journal of Climate* **31**, 4991–5014 (2018).
49. Fyfe, J. C., Saenko, O. A., Zickfeld, K., Eby, M. & Weaver, A. J. The Role of Poleward-Intensifying Winds on Southern Ocean Warming. *Journal of Climate* **20**, 5391–5400 (2007).
50. Shi, J.-R., Talley, L. D., Xie, S.-P., Liu, W. & Gille, S. T. Effects of Buoyancy and Wind Forcing on Southern Ocean Climate Change. *Journal of Climate* **33**, 10003–10020 (2020).
51. Shi, J.-R., Talley, L. D., Xie, S.-P., Peng, Q. & Liu, W. Ocean warming and accelerating Southern Ocean zonal flow. *Nature Climate Change* **11**, 1090–1097 (2021).
52. Eyring, V. *et al.* Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* **9**,

1937–1958 (2016).
53. Gillett, N. P. *et al.* The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6. *Geoscientific Model Development* **9**, 3685–3697 (2016).
54. Rayner, N. A. *et al.* Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres* **108**, (2003).
55. Huang, B. *et al.* Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons. *Journal of Climate* **30**, 8179–8205 (2017).
56. Hirahara, S., Ishii, M. & Fukuda, Y. Centennial-scale sea surface temperature analysis and its uncertainty. *Journal of Climate* **27**, 57–75 (2014).
57. Durack, P. J., Taylor, K. E., Ames, S., Po-Chedley, S. & Mauzey, C. PCMDI AMIP SST and sea-ice boundary conditions version 1.1.8. *Earth System Grid Federation* (2022) doi:https://doi.org/10.22033/ESGF/input4MIPs.16921.
58. Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C. & Wang, W. An Improved In Situ and Satellite SST Analysis for Climate. *Journal of Climate* **15**, 1609–1625 (2002).
59. Hersbach, H. *et al.* The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146**, 1999–2049 (2020).
60. Cheng, L. *et al.* Improved estimates of ocean heat content from 1960 to 2015. *Science Advances* **3**, e1601545 (2017).
61. Chan, D. & Huybers, P. Correcting Observational Biases in Sea Surface Temperature Observations Removes Anomalous Warmth during World War II. *Journal of Climate* **34**, 4585–4602 (2021).
62. Gregory, J. M. *et al.* The Flux-Anomaly-Forced Model Intercomparison Project (FAFMIP) contribution to CMIP6: investigation of sea-level and ocean climate change in response to CO2 forcing. *Geoscientific Model Development* **9**, 3993–4017 (2016).
63. Hasselmann, K. *On the Signal-To-Noise Problem in Atmospheric Response Studies*. (Royal Meteorological Society, 1979).
64. Van Den Dool, H. M., Saha, S. & Johansson, Å. Empirical orthogonal teleconnections. *Journal of Climate* **13**, 1421–1435 (2000).
65. Irving, D. B., Wijffels, S. & Church, J. A. Anthropogenic Aerosols, Greenhouse Gases, and the Uptake, Transport, and Storage of Excess Heat in the Climate System. *Geophysical Research Letters* **46**, 4894–4903 (2019).
66. Shi, J.-R., Wijffels, S. E., Kwon, Y. & Xie, S. Interhemispheric Contrasts of Ocean Heat Content Change Reveals Distinct Fingerprints of Anthropogenic Climate Forcings. *Geophysical Research Letters* **50**, (2023).
67. Allen, M. R. & Tett, S. F. B. Checking for model consistency in optimal fingerprinting. *Climate Dynamics* **15**, 419–434 (1999).
68. Bonfils, C. J. W. *et al.* Human influence on joint changes in temperature, rainfall and continental aridity. *Nature Climate Change* **10**, 726–731 (2020).

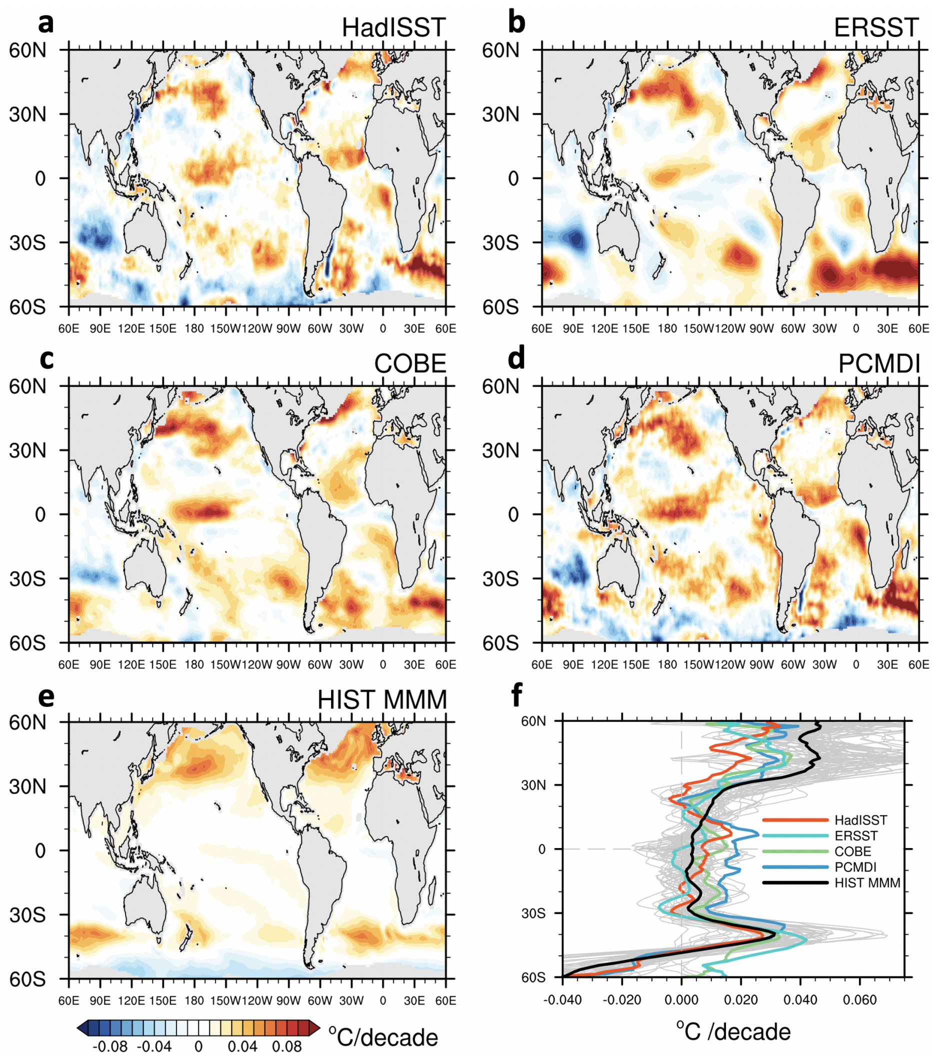


Fig. 1 Trends over 1950 to 2014 in the annual cycle amplitude of SST (SST_{AC}). a-d Trends from four observed datasets. **e** Trends from the multi-model mean (MMM) of CMIP6 HIST simulations. Regions where the model-average climatological sea-ice coverage is larger than 10% are masked in gray. **f** Zonal-mean trends in the amplitude of the SST_{AC} estimated from observations and models. The gray curves are from 51 individual HIST simulations. The domain over which all calculations are performed is restricted to 60˚S-60˚N to minimize the impact of sea ice changes on SST_{AC}.

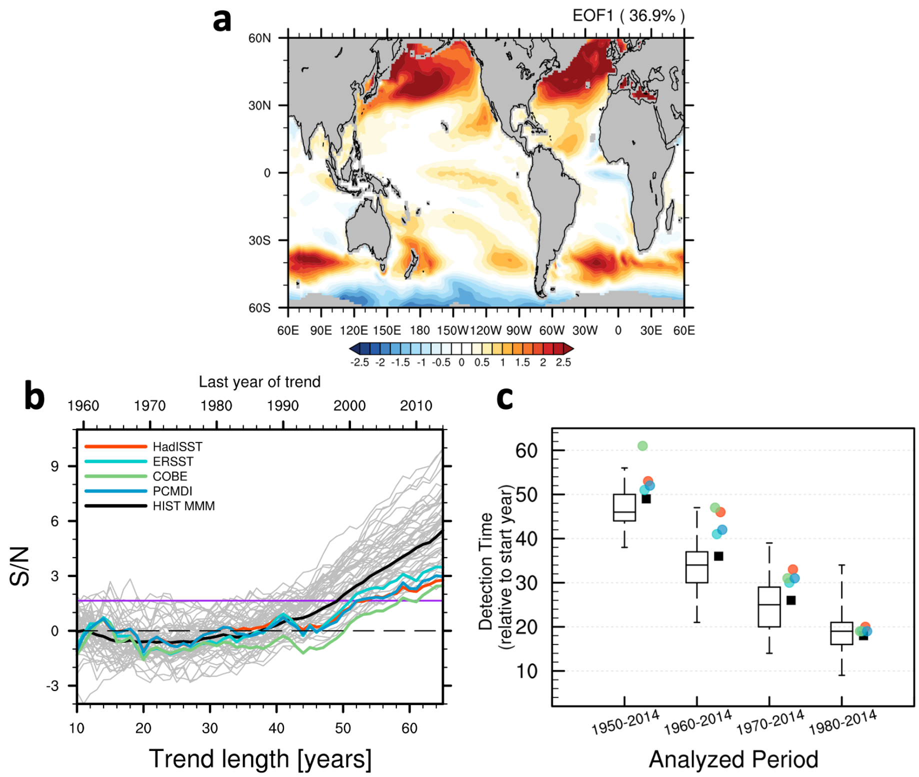


Fig. 2 SST_{AC} fingerprint and signal-to-noise (S/N) analysis. a Time-invariant HIST MMM fingerprint pattern. The fingerprint is defined here as the EOF1 of the MMM SST_{AC} changes over 1950-2014. **b** Timescale-dependent S/N ratios for trends calculated from signal and noise time series for period of 1950-2014. The HIST MMM result is the black curve; results from individual HIST runs are the gray curves. The colored lines denote S/N ratios estimated by searching for the HIST MMM SST_{AC} fingerprint in four different observed SST datasets. The horizontal purple line is the 5% significance level (see Methods). **c** Detection time relative to the start year for the model-predicted SST_{AC} fingerprint from the HIST experiment. Fingerprint detection occurs when the S/N ratios for an *L*-year analysis period first exceed the stipulated significance level and then remain above it for all larger values of *L*. The y-axis shows the value of *L* that satisfies this condition. Results are for four different assumed analysis start years (1950, 1960, 1970, and 1980). In the box-and-whisker plots, the horizontal bar is the median value, the box size represents the interquartile range, and the whiskers span the full range of detection times from all 51 individual HIST realizations. Black squares are the detection times calculated with the MMM. Colored circles are detection times estimated by searching for the model-predicted SST_{AC} fingerprint in four different observed SST datasets (see panel b legend).

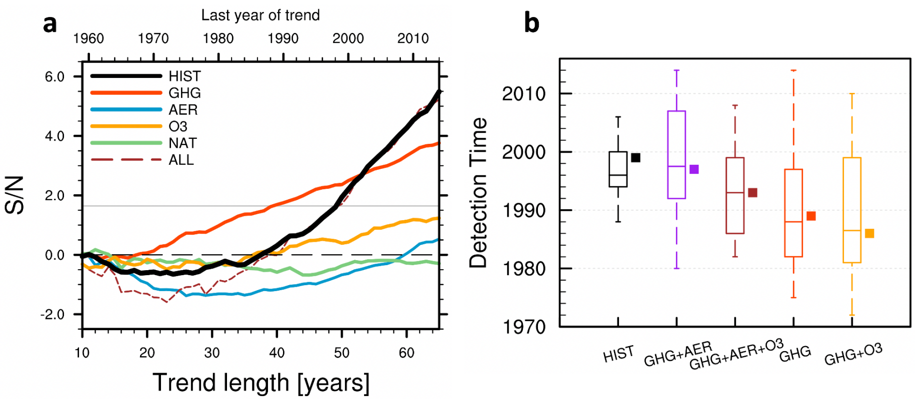


Fig. 3 S/N ratios and detection times from single-forcing runs and their linear combinations. a S/N ratios for the signal trends obtained by a fingerprint analysis involving the patterns of SST_{AC} change estimated from the MMM of different experiments. Results are for Method 1 (see Methods). For O3, the MMM is calculated from the 4 models for which O3 results were available. The MMM in the remaining cases is based on a larger set of 10 models. ALL represents the linear combination of S/N ratios from GHG, AER, O3 and NAT. The horizontal gray line is the 5% significance level. **b** The detection times of the HIST fingerprint estimated from HIST, GHG, and linear combinations of SST_{AC} changes from GHG, AER, and O3. The analysis period is 1950-2014.

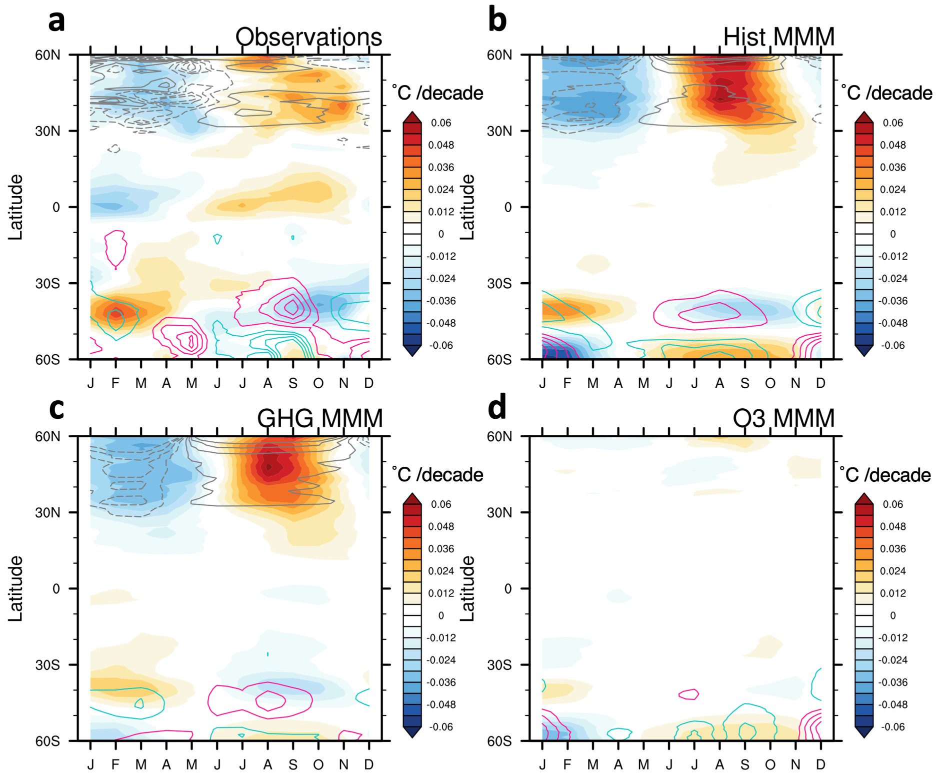


Fig. 4 Zonal-mean trends over 1950 to 2014 in monthly-mean SST, zonal wind stress, and *MLD*. **a** The ensemble mean of four different observed datasets**. b-d** The MMM of the HIST, GHG, and O3 simulations. All results are departures from annual-mean trends. Colored shading denotes monthly SST trends, gray contours are *MLD* trends plotted with a 0.75 m decade^{-1} interval, and colored contours are zonal wind stress trends plotted with a 7.2 x 10^{-4} Pa decade^{-1} interval (with positive changes shown in magenta). The zero contours are omitted. We show the *MLD* changes in the NH only and the wind stress changes in the SH only. Additional information about simulated and observed *MLD* and wind changes (including observational data sources) is given in Extended Data Fig. 9.

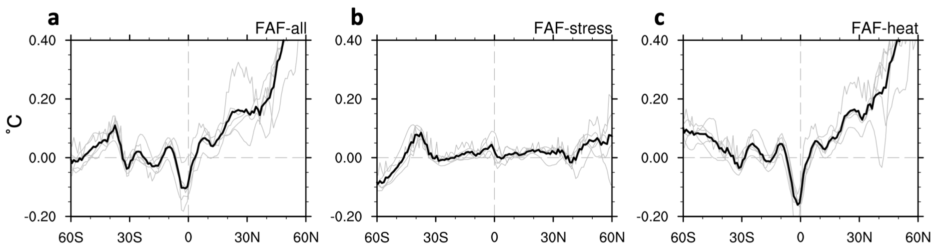


Fig. 5 Zonal-mean SST_{AC} changes from FAFMIP perturbation experiments relative to piControl. a-c Results from FAF-all, FAF-stress, and FAF-heat, respectively. For each model, the piControl results are averaged over the full length of the simulation. Perturbation results are averaged over years 31 to 70. The gray curves are from individual models and the black curves are the MMM results.

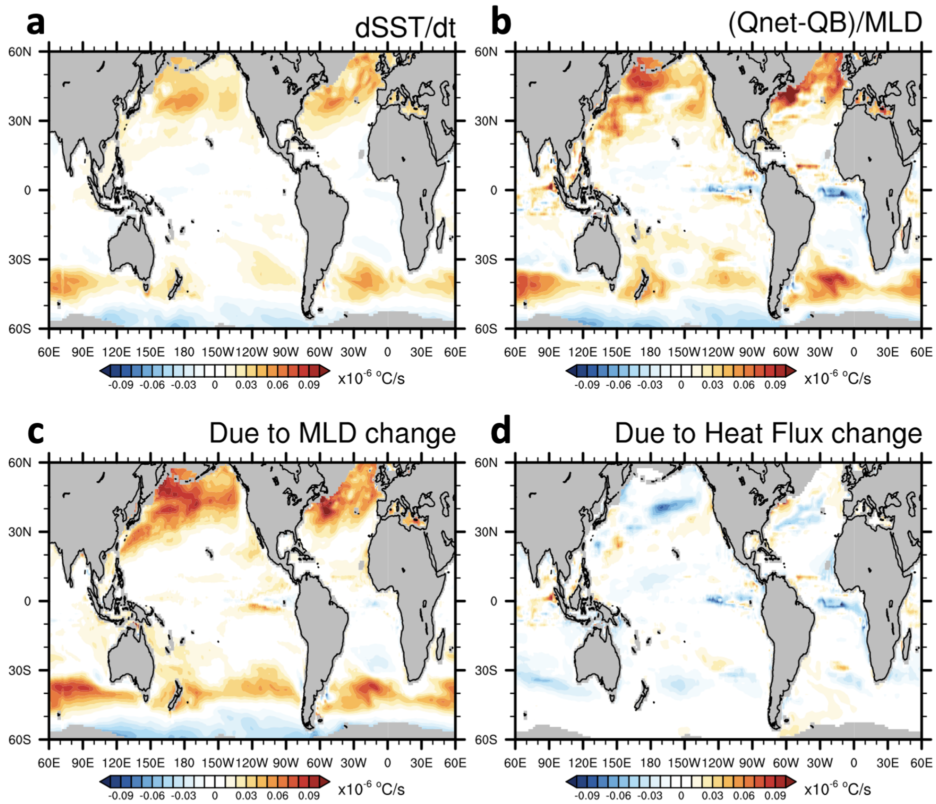
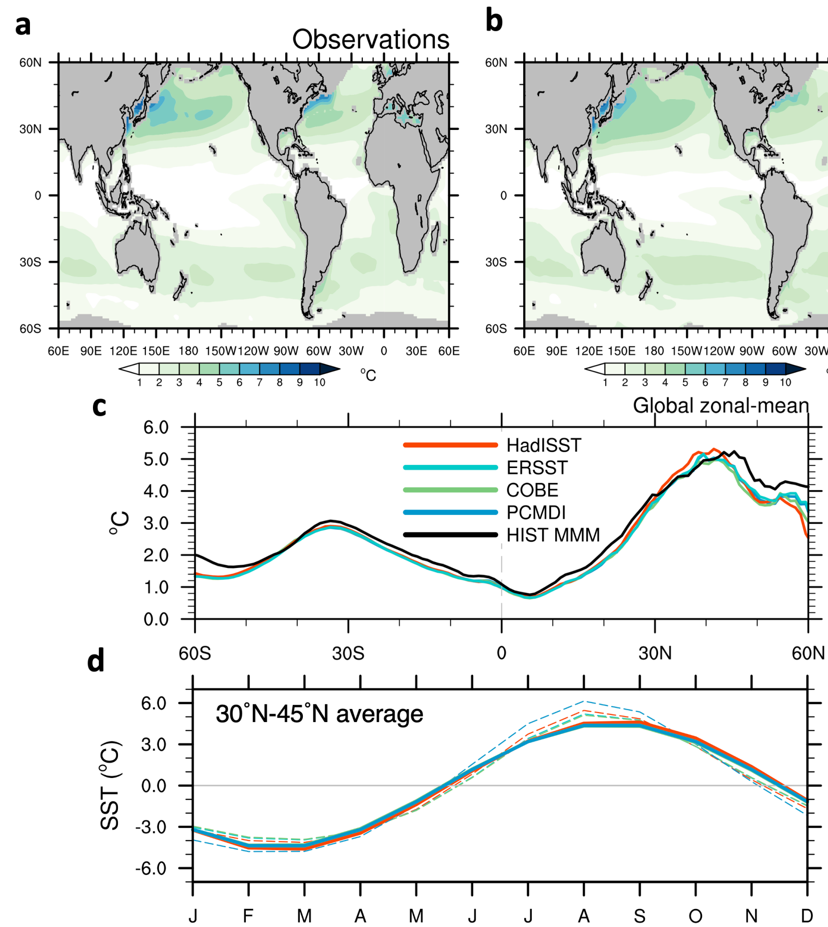
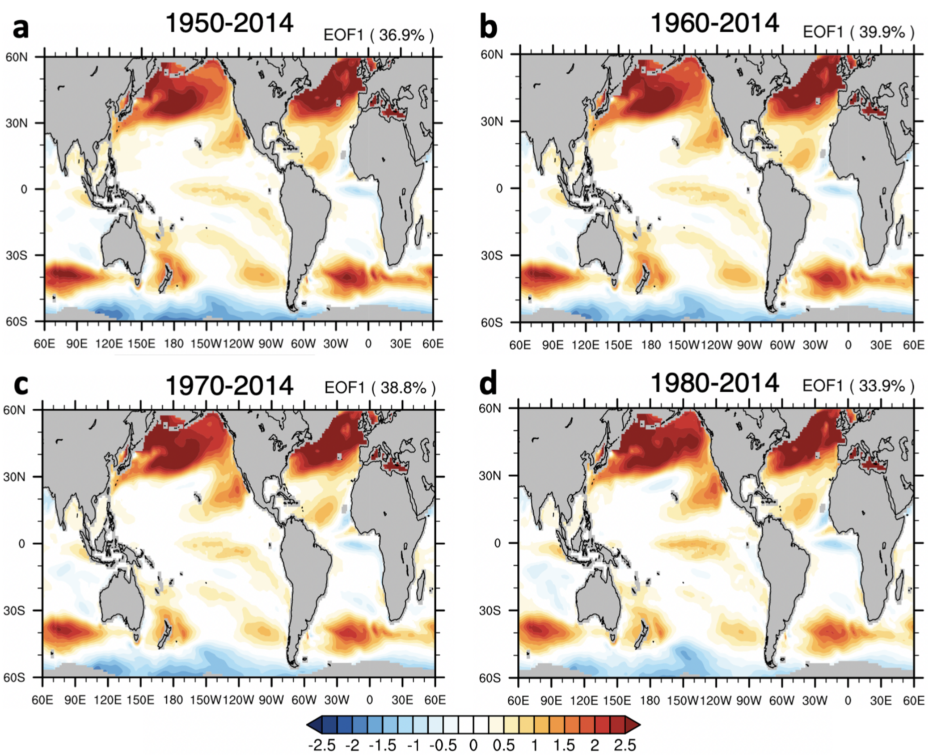


Fig. 6 Annual cycle amplitude changes between 1950-1979 and 1985-2014. a-b Changes of the annual cycle of SST tendency from CMIP6 HIST MMM and the estimate based on the mixed-layer heat budget model. **c-d** Contributions of the changing *MLD* and heat flux (*Qnet-Qb*) to the changes of the annual cycle of SST tendency, respectively.

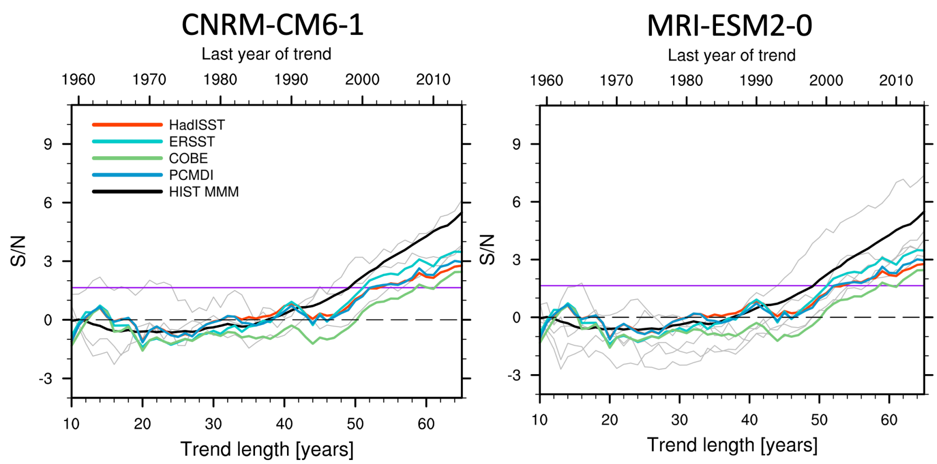
Extended Data:



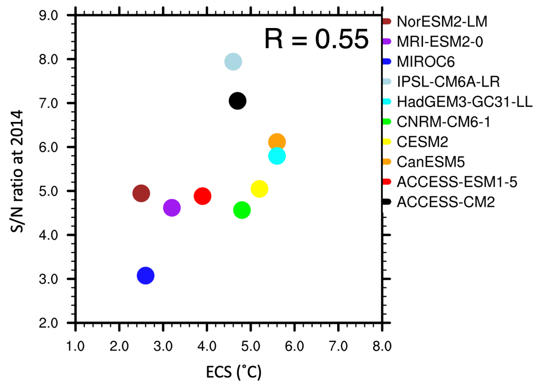
Extended Data Fig. 1. Spatial patterns and zonal mean of the climatology of SST annual cycle amplitude (SST_{AC}) from four different observational products and from the multi-model mean (MMM) of the HIST simulations. a Average of four different observed SST datasets. **b** HIST MMM. **c** Zonal-mean climatology of the HIST MMM and individual observed SST datasets. **d** Monthly climatology of SST averaged between 30˚N-45˚N from observations (dashed curves) and the fits of the first harmonic obtained through Fourier analysis (solid curves). Results are calculated over 1950 to 2014.



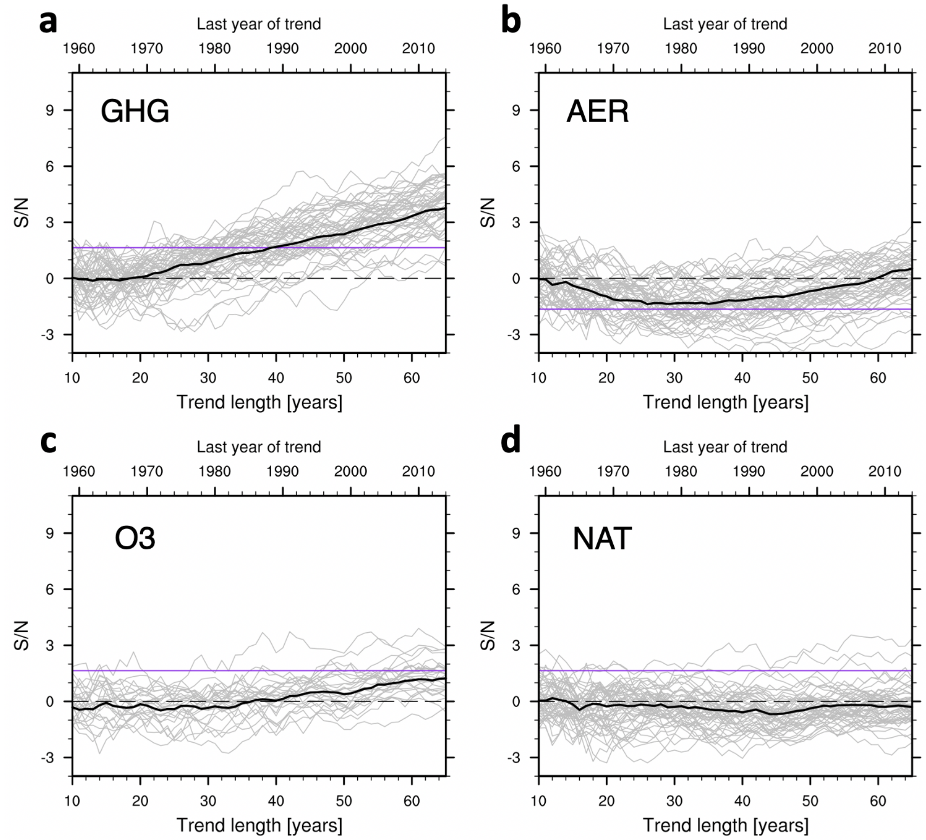
Extended Data Fig. 2. Leading EOF of SST_{AC} estimated from the HIST MMM. a-d Results for four different analysis periods. The explained variances are shown in brackets.



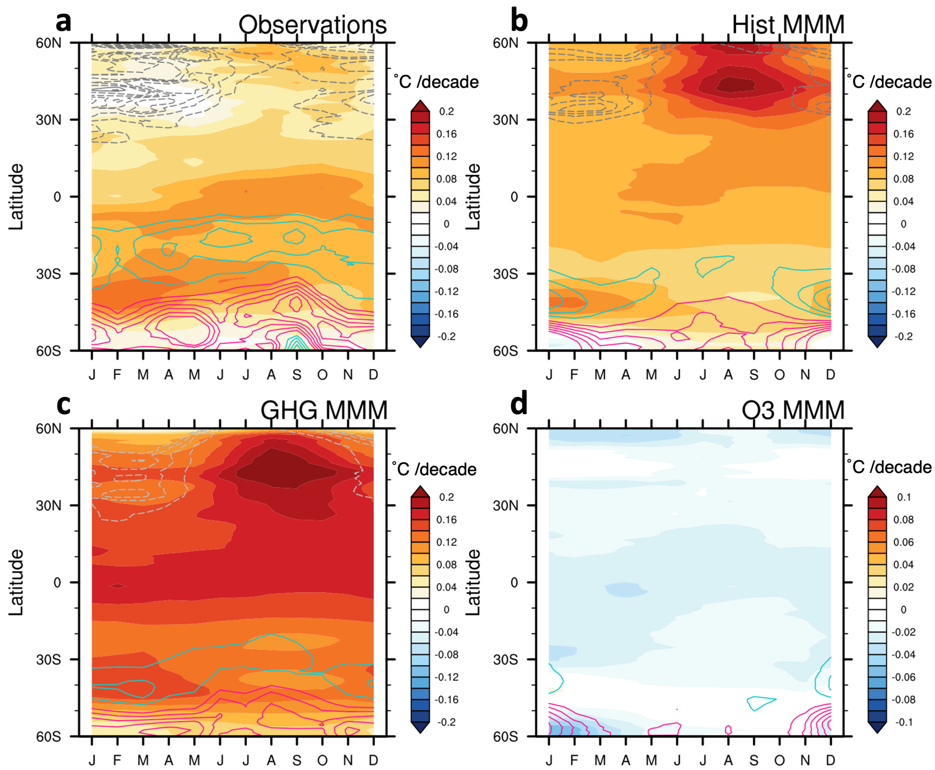
Extended Data Fig. 3. S/N ratios from two selected CMIP6 models. Results are as in Fig. 2b, but the "model only" S/N ratios here are from two models only: CNRM-CM6-1 and MRI-ESM2-0. Individual realizations from each model can have appreciable differences in their S/N behavior.



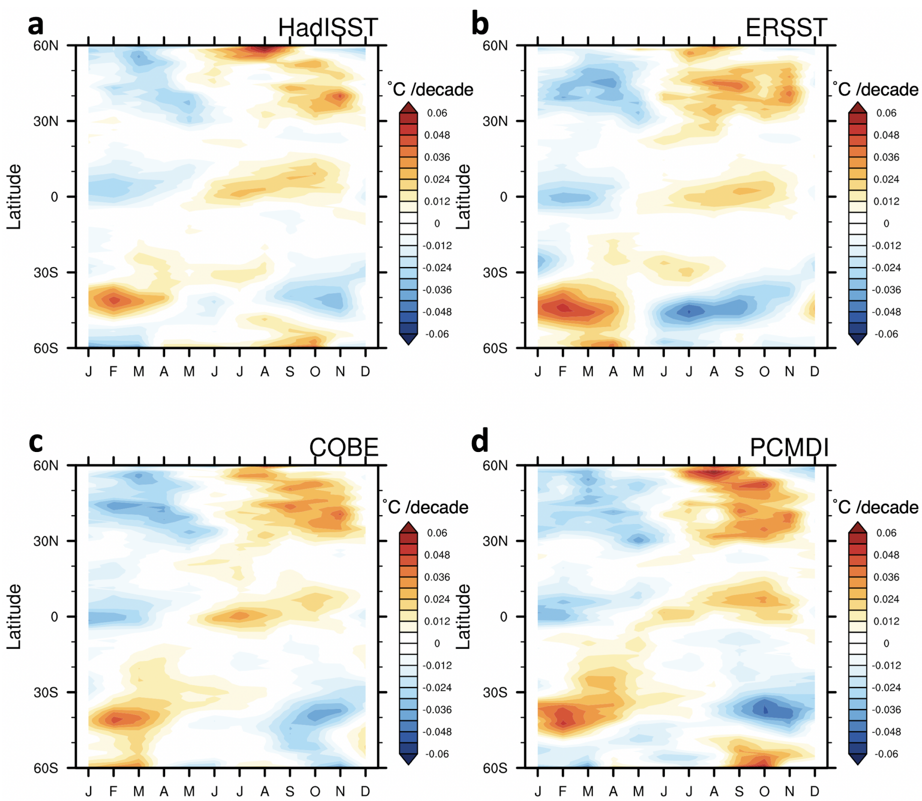
Extended Data Fig. 4. Scatterplot between the climate sensitivity of the 10 CMIP6 models analyzed here and the final value of the S/N ratio for the 65-year analysis period from 1950 to 2014. The effective climate sensitivities are based on the results from ref 40.



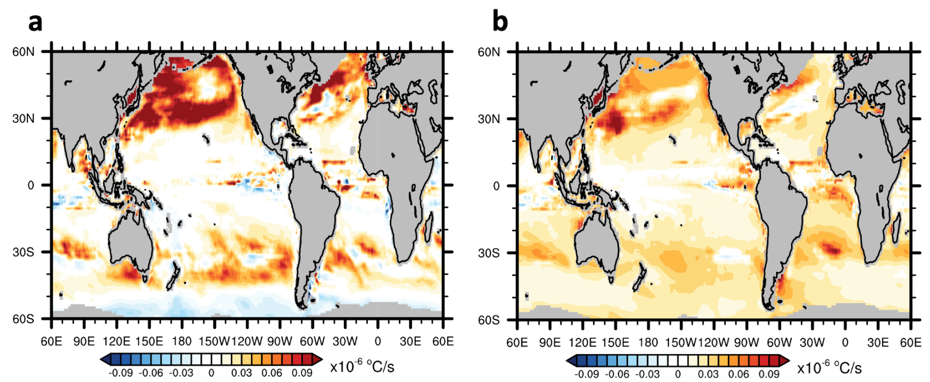
Extended Data Fig. 5. S/N ratios from the GHG, AER, O3, and NAT single-forcing runs. Results are based on use of the same HIST fingerprint, which is searched for in the SST_{AC} changes of each single-forcing run (Method 1). Each panel shows the MMM result (the black curve) and results from individual realizations (the gray curves). GHG, AER, and NAT results are from 10 models with a total of 51 realizations; only four models with a total of 26 realizations were available for calculating O3 S/N ratios. The horizontal purple line is the 5% significance level. For further details refer to Methods.



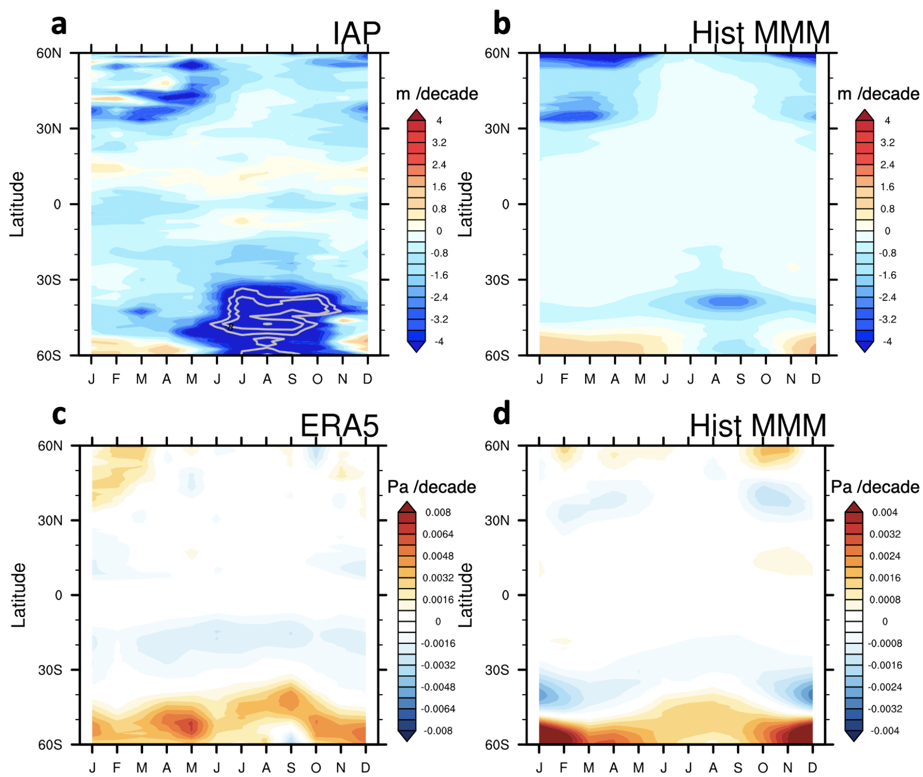
Extended Data Fig. 6. Zonal-mean monthly-mean SST trends over 1950 to 2014. a The ensemble mean of four observed datasets. **b-d** The MMM of the HIST, GHG, and O3 simulations. In contrast to Fig. 4, the trends are not expressed as departures from annual-mean trends.



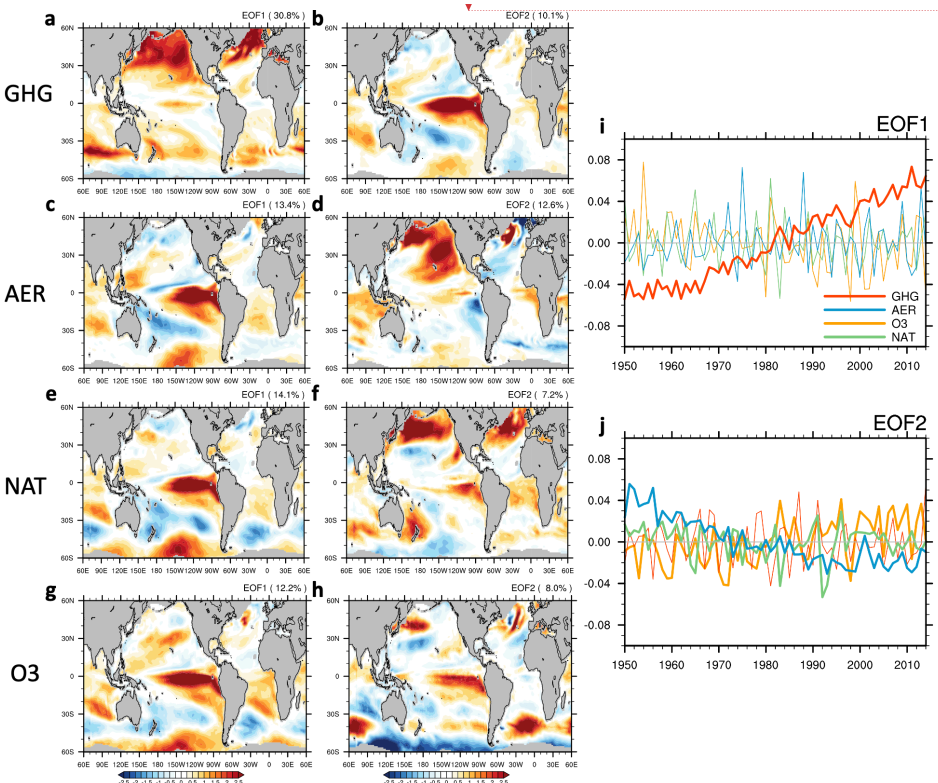
Extended Data Fig. 7. Zonal-mean monthly-mean SST trends over 1950 to 2014 in four observed datasets. The results are expressed as departures from annual-mean trends.



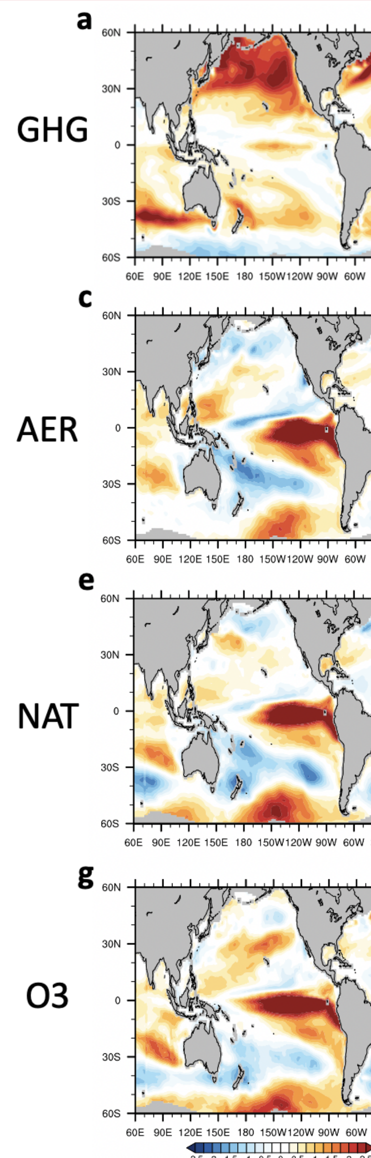
Extended Data Fig. 8. Changes of annual cycle amplitude of SST tendency between 1950-1979 and 1985-2014 due to *MLD* changes. a Changes of annual cycle when it is assumed to have a consistent summer *MLD* change for all 12 months (see Eq. 7). **b** Changes of annual cycle when *MLD* is assumed to shoal by 5% at every location and in every month relative to the background monthly value (see Eq. 8).



Extended Data Fig. 9. Zonal-mean monthly-mean trends over 1950 to 2014 in *MLD* and zonal wind stress. a-b *MLD* trends from the IAP product and the MMM of the HIST simulations, respectively. Gray contours highlight the large *MLD* trends of -6 and -8 m/decade. **c-d** Zonal wind stress trends from ERA5 and the MMM of the HIST simulations.

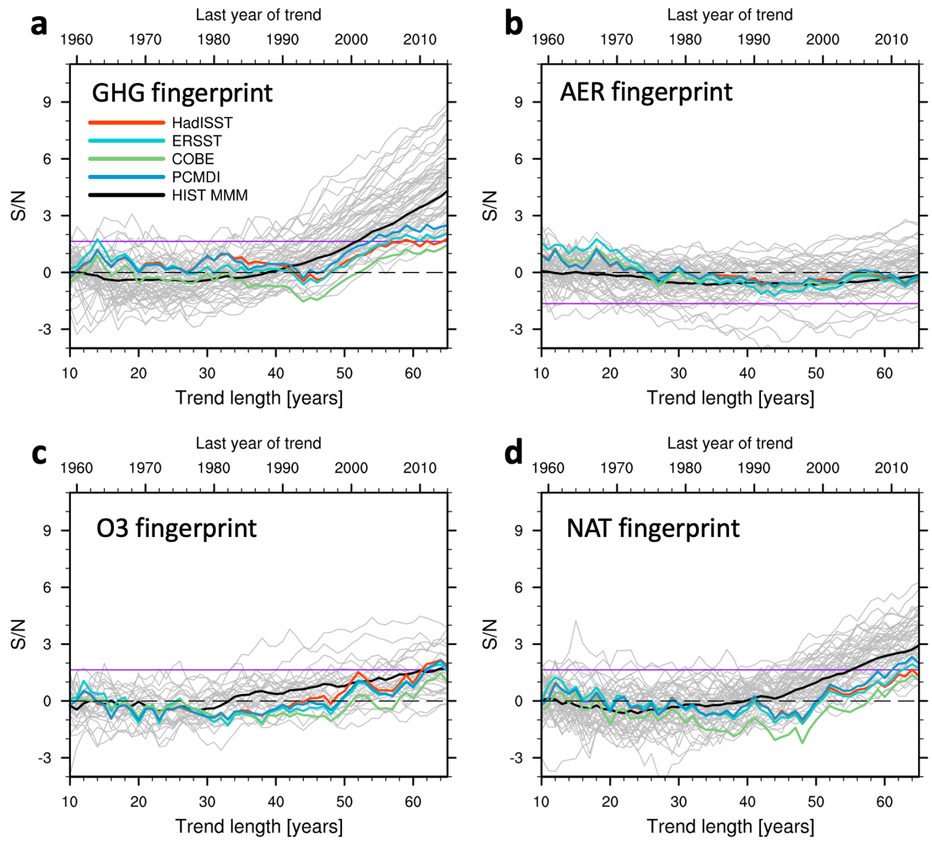


Extended Data Fig. 10. First two EOFs of SST_{AC} anomalies calculated from the MMM of the GHG, AER, NAT, and O3 single-forcing experiments. a-h Results for EOF1 and EOF2 are in the left and right columns, respectively. The explained variances are shown in brackets. i-j Principal components for EOF1 and EOF2 from four single-forcing experiments. All calculations are over 1950-2014.



Deleted:

Deleted:



Extended Data Fig. 11. S/N ratios of HIST runs and observations obtained using the fingerprints estimated from single-forcing experiments (GHG, AER, O3, and NAT). In Method 2, the SST_{AC} changes in the individual single-forcing runs are projected onto their respective fingerprints. The GHG fingerprint is the EOF1 pattern from the left column of Extended Data Fig. 10. Because the leading EOFs of AER, O3, and NAT simulations capture the effect of ENSO variability on SST_{AC}, the fingerprints for AER, O3 and NAT are the EOF2 patterns from the right column of Extended Data Fig. 10. The horizontal purple line is the 5% significance level.

Extended Data Table 1. CMIP6 models and the number of model realizations used in this study. The left column shows the 10 CMIP6 models for which HIST, GHG, AER, and NAT runs were available. For O3, results were available from four models only (right column). The middle and right columns show the number of realizations available for each model. The identifiers of these realizations (r1, etc.) are indicated in brackets.

Model names	Number of realizations used in HIST, GHG, AER, and NAT	Number of realizations used in O3
ACCESS-CM2	3 (r1-r3)	--
ACCESS-ESM1-5	3 (r1-r3)	--
CanESM5	15 (r1-15)	10 (r1-r10)
CESM2	2 (r1 and r3)	--
CNRM-CM6-1	3 (r1-r3)	--
HadGEM3-GC31-LL	4 (r1-r4)	--
IPSL-CM6A-LR	10 (r1-r3)	10 (r1-r10)
MIRO6	3 (r1-r3)	3 (r1-r3)
MRI-ESM2-0	5 (r1-r3)	3 (r1, r3, and r5)
NorESM2-LM	3 (r1-r3)	--

Extended Data Table 2. Uncentered pattern correlations between fingerprints from different experiments. For HIST and GHG, the fingerprint is obtained from the first EOF mode. For AER, O3, and NAT, the second EOF mode is used as the fingerprint (see Extended Data Fig. 10).

Pattern Correlation	HIST	GHG	AER	O3	NAT
HIST	1				
GHG	0.87	1			
AER	0.20	0.45	1		
O3	0.52	0.34	-0.06	1	
NAT	0.75	0.68	0.31	0.41	1