Exceptional Stratospheric Contribution to Human Fingerprints on Atmospheric Temperature

Benjamin D. Santer[a,b,1], **Stephen Po-Chedley**[c], **Lilong Zhao**[d], **Cheng-Zhi Zou**[e], **Qiang Fu**[f], **Susan Solomon**[g], **David W. J. Thompson**[h,i], **Carl Mears**[j], **and Karl E. Taylor**[c]

[a]Physical Oceanography Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA; [b]Joint Institute for Regional Earth System Science and Engineering, University of California at Los Angeles, Los Angeles, CA 90095, USA; [c]Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA; [d]Nanjing University of Information Science and Technology, Nanjing, China; [e]Center for Satellite Applications and Research, NOAA/NESDIS, College Park, MD 20740, USA; [f]Department of Atmospheric Sciences, University of Washington, Seattle, WA 98195, USA; [g]Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; [h]Department of Atmospheric Sciences, Colorado State University, Fort Collins, CO 80521, USA; [i]School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, UK; [j]Remote Sensing Systems, Santa Rosa, CA 95401, USA

This manuscript was compiled on March 10, 2023

In 1967, scientists used a simple climate model to predict that human-caused increases in atmospheric CO_2 should warm Earth's troposphere and cool the stratosphere. This important signature of anthropogenic climate change has been documented in weather balloon and satellite temperature measurements extending from near-surface to the lower stratosphere. Stratospheric cooling has also been confirmed in the mid- to upper stratosphere, a layer extending from roughly 25 to 50 km above Earth's surface (S_{25-50}). To date, however, S_{25-50} temperatures have not been used in pattern-based attribution studies of anthropogenic climate change. Here we perform the first such "fingerprint" study with satellite-derived patterns of temperature change that extend from the lower troposphere to the upper stratosphere. Including S_{25-50} information increases signal-to-noise ratios by a factor of five, markedly enhancing fingerprint detectability. Key features of this global-scale human fingerprint include stratospheric cooling and tropospheric warming at all latitudes, with stratospheric cooling amplifying with height. In contrast, the dominant modes of internal variability in S_{25-50} have smaller-scale temperature changes and lack uniform sign. These pronounced spatial differences between S_{25-50} signal and noise patterns are accompanied by large cooling of S_{25-50} (1-2°C over 1986 to 2022) and low S_{25-50} noise levels. Our results explain why extending "vertical fingerprinting" to the mid- to upper stratosphere yields incontrovertible evidence of human effects on the thermal structure of Earth's atmosphere.

climate change detection and attribution | stratospheric temperature | satellite data |

In simulations performed with a simple radiative convective climate model in 1967, Manabe and Wetherald progressively doubled levels of atmospheric CO_2 from 150 to 300 to 600 parts per million (1). This yielded increasing warming of the troposphere and increasing cooling of the stratosphere (2), with cooling predicted to amplify with greater height above the tropopause. The vertical profile of temperature change predicted by Manabe and Wetherald was subsequently confirmed by more complex models and by observations (3–8).

By the early 2000s, measurements of multidecadal changes in the thermal structure of the atmosphere were available from weather balloon networks (9, 10), satellite-based microwave sounders (11–13), and reanalyses (14). All three sources provided adequate spatial coverage for estimating observed patterns of zonal-mean temperature change (5–7, 15) and for comparing these patterns with vertically resolved temperature changes obtained from General Circulation Model simulations.

Early comparisons of this type noted that the observed latitude-height patterns were distinctly different from estimated patterns of natural internal variability, but consistent with the profile of atmospheric temperature change predicted by Manabe and Wetherald in response to CO_2 increases (4, 16). This early research relied on weather balloon datasets with coverage extending from the near-surface to the lower stratosphere, roughly 20 to 25 km above the surface.

Building on this pioneering work, quantitative "fingerprint" studies revealed that model-predicted latitude-height patterns of anthropogenic influence were statistically identifiable in weather balloon temperature data (15, 17). This finding has been confirmed repeatedly by subsequent investigations with newer models and improved weather balloon data sets (18, 19). The primary anthropogenic influences identified in weather balloon atmospheric temperature data are external forcings associated with increases in well-mixed greenhouse gases, the depletion and recovery of stratospheric ozone, and changes in particulate pollution (18–20).

Anthropogenic fingerprints have also been identified in atmospheric temperature measurements obtained from satellite-based Microwave Sounding Units and Advanced Microwave Sounding Units (MSU and AMSU) (21–23). As in the case

Significance Statement

Differences between tropospheric and lower stratospheric temperature trends have long been recognized as a "fingerprint" of human effects on climate. This fingerprint, however, neglected information from the mid- to upper stratosphere, 25 to 50 km above Earth's surface. Including this information improves the detectability of a human fingerprint by a factor of five. Enhanced detectability occurs because the mid- to upper stratosphere has a large cooling signal from human-caused CO_2 increases, small noise levels of natural internal variability, and differing signal and noise patterns. Extending fingerprinting to the upper stratosphere with long temperature records and improved climate models means that it is now virtually impossible for natural causes to explain satellite-measured trends in the thermal structure of Earth's atmosphere.

B.D.S., S.P.-C., Q.F., S.S., and D.W.J.T. designed the research. L.Z. and S.P.-C. calculated synthetic satellite temperatures from model simulation output. C.-Z.Z. and C.M. provided satellite temperature data and C.-Z.Z. calculated mass weights. B.D.S. performed fingerprint calculations. All authors participated in writing the paper.

The authors declare no competing interests.

[1]To whom correspondence should be addressed. E-mail: bensanter1289gmail.com

of fingerprint studies with weather balloon data, the early fingerprint work with satellite-derived atmospheric temperatures relied on data sets that did not extend higher than approximately 25 km above Earth's surface (24–27).

In consequence, all previous pattern-based studies seeking to discern a human fingerprint in weather balloon and satellite atmospheric temperature data have neglected the mid- to upper stratosphere (S_{25-50}), where the temperature signal of CO_2 increase is expected to be considerably larger than in the troposphere or the lower stratosphere (1, 8). In searching for an anthropogenic CO_2 signal, the S_{25-50} layer has the additional advantage that it is less affected than lower atmospheric layers by particulate pollution and by human-caused changes in stratospheric ozone (28).

Satellite-based Stratospheric Sounding Units (SSU) provide temperature changes for the S_{25-50} layer (29). Initial SSU-based temperature-change estimates obtained by two different groups diverged markedly (8) but are now in closer agreement (27, 30, 31).[*] Only one group, however, supplies spatially resolved SSU data suitable for pattern-based fingerprint studies and has merged SSU data with AMSU-A data (AMSU-A also samples the S_{25-50} layer). Merging allows extension of SSU data beyond 2006 (27), yielding a continuous record of mid- to upper stratospheric temperature change from 1986 to the present.[†] We refer to this merged product as "SSU". Merged MSU and AMSU data, which sample the troposphere and lower stratosphere, are referred to as "MSU".

Here we expand upon earlier fingerprint studies that relied solely on MSU data for estimating latitude-height profiles of atmospheric temperature change (23). We leverage the availability of improved SSU and MSU data sets and of newer simulations (32) performed with models with higher tops, which permits calculation of synthetic SSU temperatures from simulation output. We analyze atmospheric temperature signals from a multi-model ensemble of historical simulations (HIST_{ext}) that have been extended after 2014 with results from a specific climate change scenario. We also rely on an ensemble of pre-industrial control runs with no year-to-year changes in human or natural external factors. The control runs provide multi-model estimates of the "noise" of natural internal variability. Model signal and noise estimates are essential ingredients of fingerprint studies (23, 33, 34).

It is not obvious *a priori* how incorporating the mid- to upper stratosphere will affect signal-to-noise (S/N) ratios and the detectability of an anthropogenic fingerprint. While model and observed cooling signals in S_{25-50} are $\approx$ 1-2°C over the satellite era (8, 31, 35), the noise of natural internal variability can be appreciable on monthly timescales, partly due to the impact of sudden stratospheric warming events on S_{25-50} temperatures over the Arctic (36). Additionally, it must be determined whether human-caused signals and natural variability have similar temperature-change patterns in the S_{25-50} layer – a situation which would be unfavorable for signal identification (37). Although previous investigations have compared simulated and observed global-mean temperature changes in the S_{25-50} layer (8, 31, 35), our study is the first to perform pattern-based fingerprinting with temperature changes extending from the lower troposphere to the upper stratosphere.

[*] This agreement does not necessarily signify that observational uncertainties in SSU data are trivially small. The process of identifying and adjusting for complex non-climatic factors is ongoing and benefits from the involvement of multiple independent scientific groups.

[†] The SSU record commences in 1979, but several SSU channels have data gaps prior to 1986 (29).

We rely on satellite data from three groups and on model data from phase 6 of the Coupled Model Intercomparison Project (CMIP6) (32). Our focus is on temperature changes in six atmospheric layers: SSU channels 3, 2, and 1 and MSU retrievals for the lower stratosphere (TLS), the total troposphere (TTT), and the lower troposphere (TLT). The approximate peaks of the weighting functions for these six layers are at 45, 38, 30, 19, 5.6, and 3.1 km above Earth's surface (respectively). Further details of all data sets and analysis methods are given in the Materials and Methods and the Supporting Information (SI) Appendix.

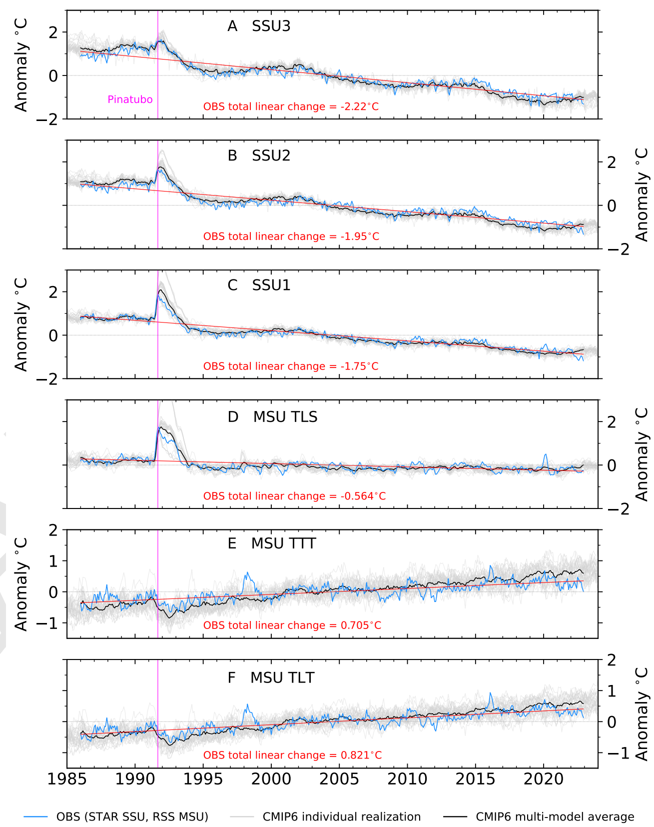

Fig. 1. Observed and simulated changes in global-mean monthly-mean temperature in six atmospheric layers. Results are temperatures from channels 3, 2, and 1 of the Stratospheric Sounding Unit (SSU; panels A-C) (27), lower stratospheric temperature from the Microwave Sounding Unit (MSU TLS; panel D), MSU total tropospheric temperature (TTT; panel E) and MSU lower tropospheric temperature (TLT; panel F) (25). The peaks of the weighting functions for these six layers are at *ca.* 45, 38, 30, 19, 5.6, and 3.1 km above Earth's surface (respectively). Results are anomalies relative to climatological monthly means over 1986 to 2022. Model simulations are from nine different CMIP6 models and a total of 32 realizations of historical climate change (see Methods and SI).

Global-mean changes

Consistent with the early Manabe and Wetherald predictions of the atmospheric temperature response to CO_2 increase (1), both the satellite data and simulations performed with state-of-the-art Earth System Models (ESMs) show tropospheric warming and stratospheric cooling over 1986 to 2022 (Fig. 1) (31, 35, 38, 39). Other common features in models and satellite data include amplification of cooling with increasing height in the stratosphere (8, 31, 35), short-term stratospheric warming after the 1991 Pinatubo eruption (with warming decreasing in amplitude with increasing stratospheric height), longer-term tropospheric cooling following Pinatubo (40), and a roughly 11-year solar signal in the SSU channels (8, 35).

Noticeable model-versus-observed differences include overestimated model-average stratospheric cooling and larger model-average tropospheric warming trends (Fig. 2). The latter discrepancy is due to multiple factors, including model-versus-observed differences in the phasing of multidecadal Pacific internal variability (41), model forcing and response errors (42–44), and the relatively limited ensemble size of HIST$_{ext}$ runs available here (41). Residual errors in observed satellite data are also a possible contributory factor (39).

In the three SSU channels, the stratospheric cooling trends over 1986 to 2022 in satellite data and HIST$_{ext}$ runs are over an order of magnitude larger than control run estimates of the natural internal variability of 37-year atmospheric temperature trends (Figs. 2A-C). The amplitudes of forced and unforced trends are more similar in the lower stratosphere and troposphere, although all satellite and HIST$_{ext}$ TLS, TTT, and TLT trends are still clearly separated from their respective control run distributions (Figs. 2D-F). These results indicate that at the global-mean level, the S/N properties of the S$_{25-50}$ layer are highly favorable for anthropogenic signal detection.

The analysis in Fig. 2 is over 1986 to 2022 only – the period of continuous coverage of SSU and MSU temperature measurements. This period samples both the pronounced depletion of stratospheric ozone in the last three decades of the 20th century and the gradual recovery of stratospheric ozone in the early 21st century (28, 45). In addition to ozone, other atmospheric constituents can also show important time variations in radiative forcing (46–49). It is of interest here to consider the impact of such variations on simulated temperature-change profiles, and to explore how S/N properties changes as the net anthropogenic forcing changes.

Figure 3 shows simulated global-mean temperature changes in the HIST$_{ext}$ runs. Results are for four different 25-year time windows: 1950-1974, 1975-1999, 2000-2024, and 2025-2049. The second and third periods sample times influenced by ozone depletion and ozone recovery (respectively) (28, 45); the fourth period has substantially larger net anthropogenic forcing than the first. As in Fig. 2, control run trend distributions provide information on the magnitude of unforced atmospheric temperature changes. This information is valuable for assessing the significance of the forced temperature trends in the HIST$_{ext}$ simulations.

Consider the troposphere first. In TLT and TTT, each successive 25-year period has larger ensemble-mean tropospheric warming and greater separation from the mean of the sampling distribution of unforced trends (i.e., higher S/N levels). This progressive warming is consistent with increasing positive forcing by anthropogenic greenhouse gases. The early 1950-1974 period has large, time-increasing negative anthropogenic sulfate aerosol forcing (49), which helps to explain why the ensemble-mean HIST$_{ext}$ tropospheric temperature trends over this period are close to zero. Anthropogenic sulfate aerosol forcing decreases nonlinearly in the three subsequent analysis periods (49, 50), yielding a decrease in sulfate aerosol-induced tropospheric cooling. Although these pronounced temporal changes in anthropogenic sulfate aerosol forcing influence TLT and TTT, they have minimal effect on simulated stratospheric temperature trends.

In the three SSU channels, stratospheric cooling occurs in each of the four analysis periods and in every HIST$_{ext}$ realization (Figs. 3A-D). As in the case of the 1986-2022 period, cooling in the HIST$_{ext}$ runs amplifies with increasing height and is invariably significantly larger than 25-year trends arising from internal variability. One key difference relative to the tropospheric results in Figs. 3E,F is that stratospheric cooling does not increase monotonically as the 25-year analysis window advances. The effect of the large stratospheric ozone depletion over 1975-1999 is to augment CO_2-induced stratospheric cooling. As a result, the ensemble-mean HIST$_{ext}$ cooling of each SSU channel (and of TLS) is larger over 1975-1999 than in the subsequent 2000-2024 period. By 2025-2049, the primarily CO_2-driven cooling of the S$_{25-50}$ layer exceeds the CO_2 and ozone-driven S$_{25-50}$ cooling over 1975-1999.

Figure 3 shows that despite important changes over time in the relative contributions of ozone and GHG forcing, the simulated global-mean temperature change profile in response to anthropogenic forcing is remarkably robust over 1950 to 2049. The temperature-change contrasts between tropospheric warming and cooling of the mid- to upper stratosphere generally increase with time and with larger net anthropogenic forcing and become easier to discriminate from natural internal variability. The exception is in the lower stratosphere, where forced temperature changes become less significant in the second half of the 21st century. This is due to two factors. First, lower stratospheric cooling due to GHG increases is partly offset by warming arising from the recovery of stratospheric ozone (28, 45). Second, the TLS weighting function receives a small contribution from CO_2-induced warming of the tropical upper troposphere (51). As tropical upper tropospheric warming increases over time (and as the height of the tropical tropopause increases), this contribution becomes larger.

Latitude-height trend patterns

Latitude-height patterns of atmospheric temperature trends are shown in Figs. 4A-L. In all nine models and in observations, tropospheric warming is hemispherically asymmetric, with larger warming over the Arctic than over the Antarctic. This asymmetry has multiple causes, including reduction in atmospheric burdens of anthropogenic aerosols, feedbacks associated with substantial changes in Arctic sea ice extent over the satellite era (52, 53), and hemispheric differences in ocean circulation and heat uptake (54).

In satellite data, stratospheric cooling over 1986 to 2022 is also asymmetric, with larger cooling over the Arctic and upward extension of a reduced cooling signal over the Antarctic (Figs. 4K, L). Some models capture aspects of this upward extension at mid- to high southern latitudes (Figs. 4B, C, F, G, H, and I), but most models lack the observed south-to-north decrease in S$_{25-50}$ and the maximum Arctic cooling in S$_{25-50}$.

The observed global-scale cooling of the S$_{25-50}$ layer is noticeably larger over 1986 to 2000 than over 2001 to 2022 (SI Figs. S1A, B). Larger stratospheric cooling in the earlier period is partly due to recovery from Pinatubo-induced stratospheric warming (Figs. 1A-D). The CMIP6 multi-model average captures time-evolving behavior similar to that in the satellite data, but lacks the prominent observed Arctic cooling of S$_{25-50}$ over 1986 to 2000 (SI Figs. S1C, D). As in the case of model-versus-observed stratospheric cooling differences over the longer 1986 to 2022 period, this discrepancy over the Arctic is likely related to multiple factors (see Conclusions).

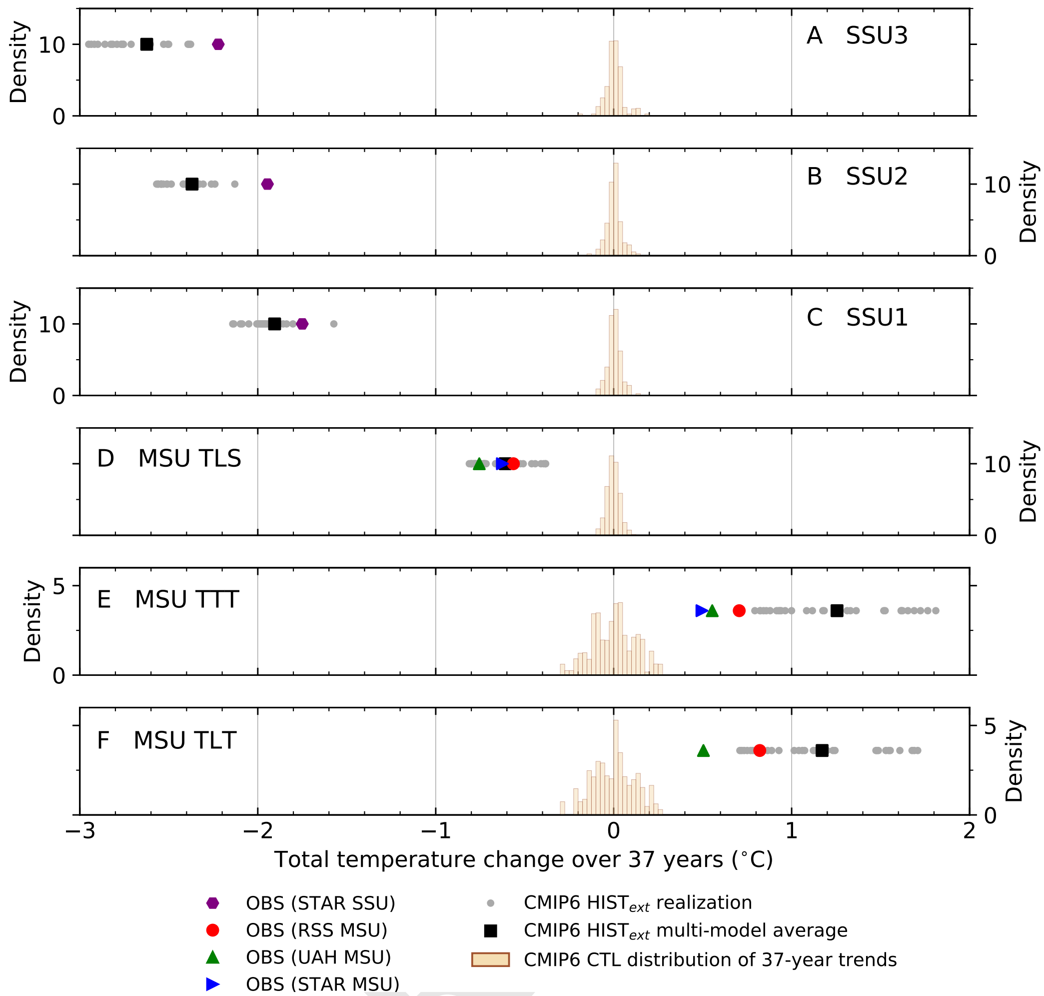

Fig. 2. Total global-mean atmospheric temperature changes over 37-year periods. Results are for six different atmospheric layers, arranged vertically by height of the layer with respect to Earth's surface (panels A-F). The total temperature change is the least-squares linear trend per year $\times$ 37 years, calculated over 1986 to 2022 for the HIST$_{ext}$ realizations and satellite observations and over 37-year non-overlapping segments of pre-industrial control runs. The latter provide estimates of the natural internal variability of atmospheric temperature trends inferred from nine different CMIP6 models. The same nine models were used to calculate the multi-model average synthetic SSU and MSU atmospheric temperature trends from 32 realizations of HIST$_{ext}$ runs with anthropogenic and natural external forcing. Trends from individual HIST$_{ext}$ realizations are also shown. See SI for details of control run trend distributions and sources of observed data. The y-axis location of the HIST$_{ext}$ trends and observed trends is arbitrary.

Fingerprint results

We use a standard pattern-based fingerprint method (23, 33, 55). This yields S/N ratios as a function of L, the timescale in years. The fingerprint F is estimated from the multi-model average latitude-height temperature changes in the HIST$_{ext}$ simulations. The signal S is a measure of the pattern similarity between F and time-varying patterns of temperature changes in observations or in individual HIST$_{ext}$ simulations. The noise N provides information on the similarity between F and time-varying patterns of natural internal variability in model control runs (see Methods and SI). If S/N ratios are larger than 3, it is unlikely that the time-increasing similarity between F and the satellite data could be due to internal variability alone (55).

Since our interest is in exploring the impact of temperature changes from different atmospheric layers on S/N properties, we show the signals calculated with fingerprints for four different spatial domains (Fig. 5A). We refer to these domains subsequently as TROP, SSU, MSU, and SSU+MSU. They comprise the two tropospheric layers in Fig. 1, the three SSU channels, the three MSU layers, and all six layers (respectively). Fingerprints estimated from the multi-model average atmospheric temperature changes for these four domains are shown in the left column of Fig. 6. The fingerprints are dominated by anthropogenic external forcing (see SI).

Consistent with the size of the global-mean temperature changes in Fig. 2, the largest signals in Fig. 5A are for the two domains (SSU and SSU+MSU) that include the large temperature changes in the mid- to upper stratosphere; the smallest signals are for MSU and TROP. This ordering of signal strength holds for the simulations and for the observations. The model spread in $S(L)$ is greater for smaller values of L, reflecting the larger noise of internal variability on shorter timescales (56). On longer multidecadal timescales, the main drivers of spread in $S(L)$ are inter-model forcing and response differences (57).

Values of $S(L)$ decrease for analysis periods ending in 1991, gradually recovering over the following 4-5 years (Fig. 5A). This decrease in $S(L)$ is due to the short-term stratospheric

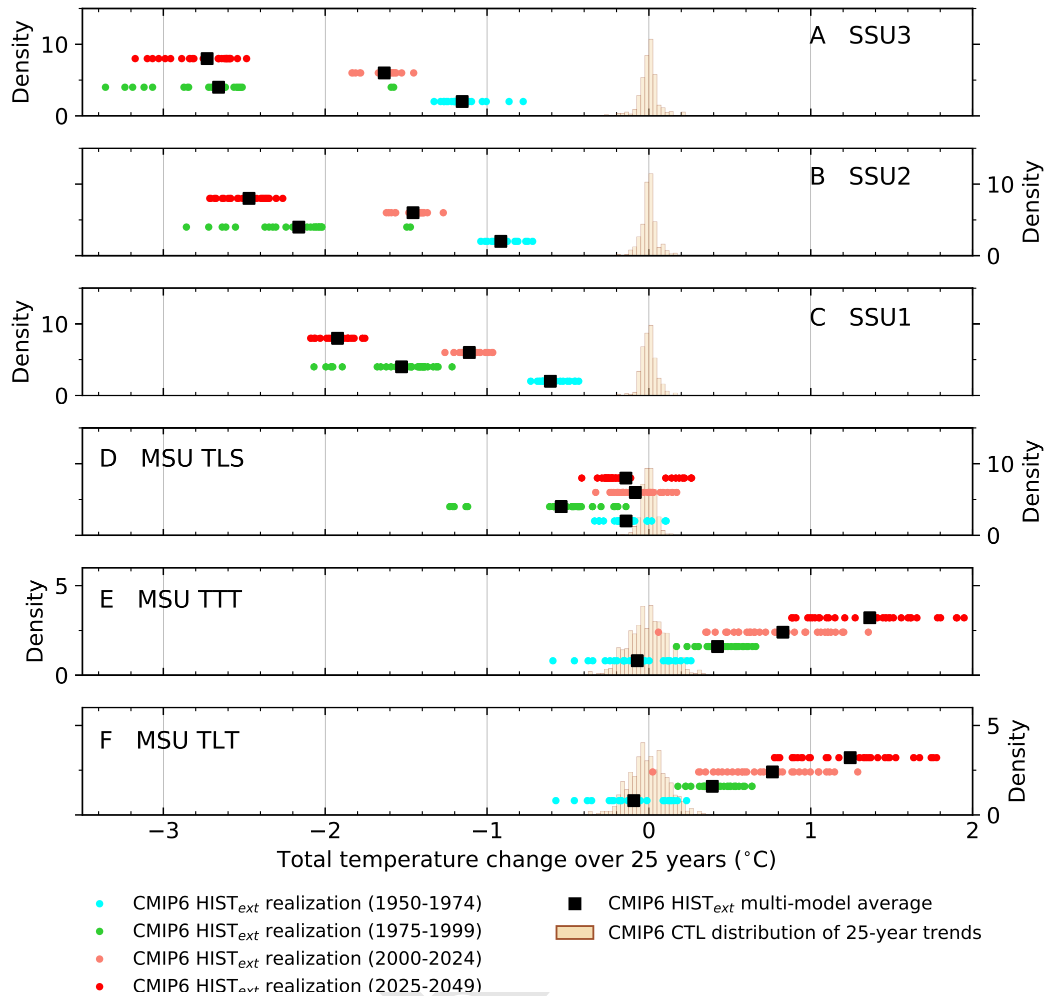

Fig. 3. Sensitivity of global-mean atmospheric temperature changes to the choice of analysis period. The total temperature change is the least-squares linear trend per year × 25 years, calculated over four different periods for the HIST$_{ext}$ realizations (1950-1974, 1975-1999, 2000-2024, and 2025-2049) and over 25-year non-overlapping segments of CMIP6 pre-industrial control runs. See Fig. 2 for analysis details and the SI for details of control run trend distributions. The y-axis location of the HIST$_{ext}$ trends is arbitrary.

warming and tropospheric cooling caused by the 1991 Pinatubo eruption – temperature changes that are of opposite sign to the searched-for anthropogenic fingerprints (see Figs. 6A, D, G, and J). For the SSU+MSU and SSU domains, stratospheric cooling during the recovery from the Pinatubo eruption augments the gradual anthropogenically induced stratospheric cooling and produces a rapid increase in signal strength over 1992 to 1997.

For all four atmospheric regions, the noise N decreases as L increases (Fig. 5B). Values of N are largest for TROP and MSU and smallest for SSU+MSU and SSU – the reverse of the ordering for signal strength in Fig. 5A. Dividing $S(L)$ by the respective value of $N(L)$ yields the signal-to-noise ratio $\mathrm{SN}(L)$ in Fig. 5C. This ratio is markedly smaller for TROP and MSU than for SSU and SSU+MSU. In the three satellite data sets, $\mathrm{SN}(L)$ for the 37-year signal trend over the 1986 to 2022 period varies between 4.6 and 6.6 for TROP, 6.7 and 9.0 for MSU, and 37.3 and 38.7 for SSU+MSU. For the SSU domain, $\mathrm{SN}(L)$ over the full analysis period is 49.3 in the only available satellite data set (27). In all four latitude-height domains, the model-predicted fingerprints in Fig. 6 are identifiable with high statistical confidence (at or above the 1% level) in each of the 32 HIST$_{ext}$ realizations and in each of the three observed data sets.[‡]

One of the key inferences from Fig. 5C – and a central finding from our study – is that extending vertical fingerprinting from "MSU space" to combined "SSU+MSU space" amplifies signal-to-noise ratios in satellite data by a factor of ≈ 5 for $\mathrm{SN}(L)$ calculated over the full 1986 to 2022 period. The inclusion of temperature changes in S_{25-50} is therefore useful in discriminating between anthropogenically driven atmospheric temperature change and internally generated variability. This enhancement of $\mathrm{SN}(L)$ in SSU+MSU is partly due to the large amplitude of the signal and the relatively low noise amplitude in S_{25-50} (Fig. 2). Signal-to-noise enhancement also reflects relative differences in the spatial similarity between the fingerprint F and the leading patterns of natural internal variability in the SSU and MSU domains (see below).

Patterns of signal and noise modes

In the fingerprint for each of the four domains considered here, temperature changes for individual satellite sounding channels

[‡]We note, however, that the three observational data sets are not independent for the SSU or SSU+MSU domains – all share the same STAR SSU data.

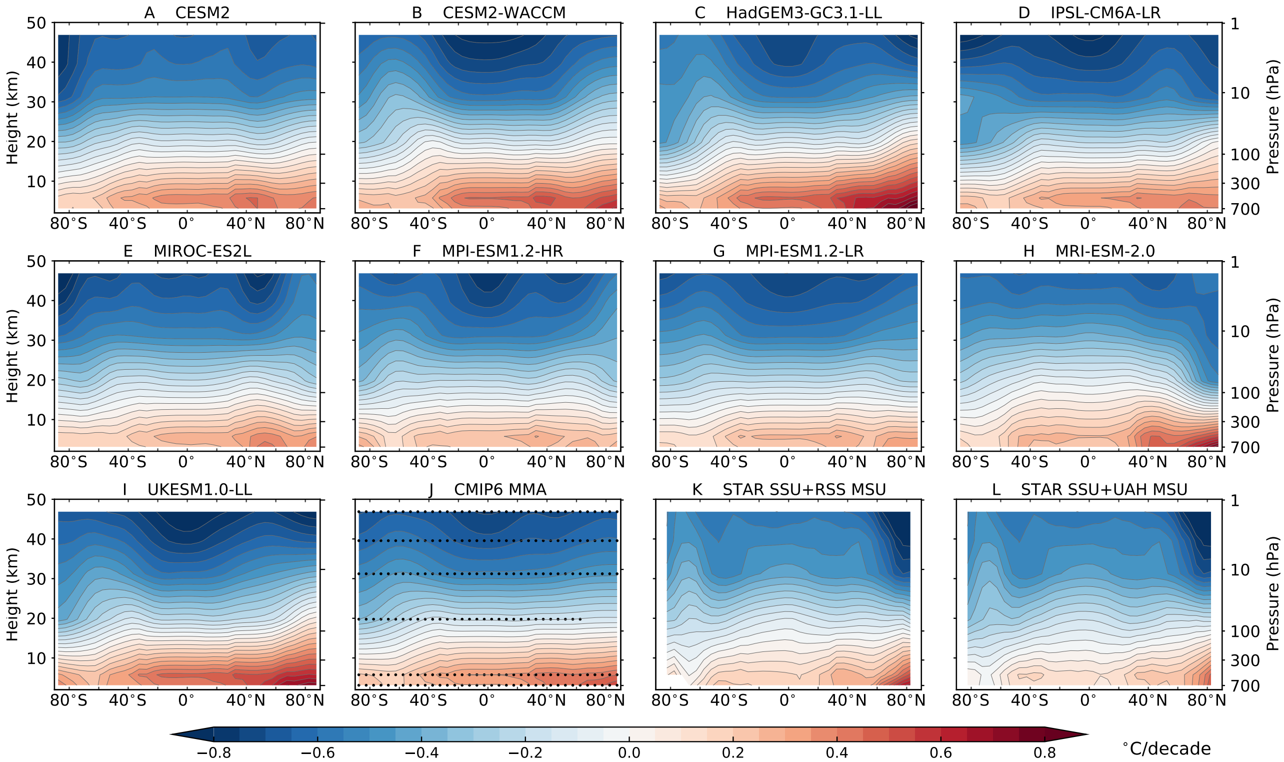

Fig. 4. Simulated and observed latitude-height profiles of atmospheric temperature trends over 1986 to 2022 (in °C/decade). Trends were calculated from zonal-mean temperatures for the six atmospheric layers in Fig. 1. Trends are plotted at the approximate heights of the maxima of each weighting function peak and were smoothly interpolated in the vertical. Model results are for HIST$_{ext}$ simulations performed with nine different CMIP6 models (panels A-I). If more than one HIST$_{ext}$ realization was available for an individual model, the result in panels A-I is for the ensemble-mean trends. The CMIP6 multi-model average is also shown (MMA; panel J). Satellite observations are for SSU data combined with two different observed MSU data sets (panels K and L; see Methods). Stippling in panel J denotes latitude bands and layers at which the local S/N ratio exceeds 2 – i.e., where the CMIP6 MMA trend is two times greater than the between-model standard deviation of the trend. The stippling indicates that at each latitude and for each of the six atmospheric layers, the MMA temperature trends are large relative to the between-model standard deviation of trends. The sole exception is in TLS over the Arctic, where there are noticeable inter-model trend differences.

vary with latitude but remain either positive or negative across all latitudes (see left column of Fig. 6). In terms of vertical structure, the fingerprints for the MSU and SSU+MSU domains are characterized by a reversal with height in the sign of temperature change (Figs. 6D and J), consistent with the large tropospheric warming and stratospheric cooling signals common to the models analyzed here (see Figs. 4A-I). Other prominent fingerprint features include Arctic amplification of low-latitude warming in TROP and amplification of stratospheric cooling with increasing height in the SSU domain (Figs. 6A and G, respectively).

In contrast to the fingerprint patterns, the leading multi-model noise modes in the middle and right columns of Fig. 6 display smaller-scale variability with pronounced meridional structure. For a given sounding channel, no noise mode has temperature changes with uniform sign at all latitudes. In the TROP domain, the leading noise mode reveals internal variability that is anticorrelated between the tropics and midlatitudes (Fig. 6B). This behavior is consistent with temperature fluctuations associated with the El Niño/Southern Oscillation (ENSO) (41). For the SSU domain, the variability in the leading noise mode is strongly anticorrelated between the tropics and extratropics (Fig. 6H), likely due to tropical upwelling and polar downwelling driven by the shallow branch of the Brewer-Dobson circulation (BDC). The noise modes for the MSU and SSU+MSU domains capture aspects of both ENSO- and BDC-induced internal variability.

To quantify the spatial similarity between fingerprint and noise patterns in Fig. 6, we calculated $r\{\mathrm{F{:}N1}\}$ and $r\{\mathrm{F{:}N2}\}$, the uncentered pattern correlations between F and the first two noise modes of the concatenated control runs (37). Values of $r\{\mathrm{F{:}N1}\}$ and $r\{\mathrm{F{:}N2}\}$ are smallest for the SSU and SSU+MSU domains and largest for TROP and MSU (see SI Fig. S2). This difference in pattern similarity across domains holds for fingerprints calculated from individual CMIP6 HIST$_{ext}$ realizations, the HIST$_{ext}$ multi-model average, and satellite data sets. The small $r\{\mathrm{F{:}N1}\}$ and $r\{\mathrm{F{:}N2}\}$ values for the SSU and SSU+MSU domains help to explain their large S/N ratios in Fig. 5C – the fingerprints for these two domains are more effective in filtering out internal variability noise.

For individual spatial domains, the clustering of points with similar correlation values in SI Fig. S2 implies that the fingerprints estimated from individual model results or individual observational data sets are spatially similar. We show this fingerprint similarity for the specific case of the SSU+MSU domain (SI Fig. S3). The fingerprints in SI Fig. S3 are the leading Empirical Orthogonal Function (EOF) of the individual model HIST$_{ext}$ simulations and the satellite data sets.

It is likely that higher-order EOFs capture additional forced components of atmospheric temperature change, such as the nonlinear TLS response to time-evolving forcing by lower stratospheric ozone depletion (28, 58). This is illustrated by

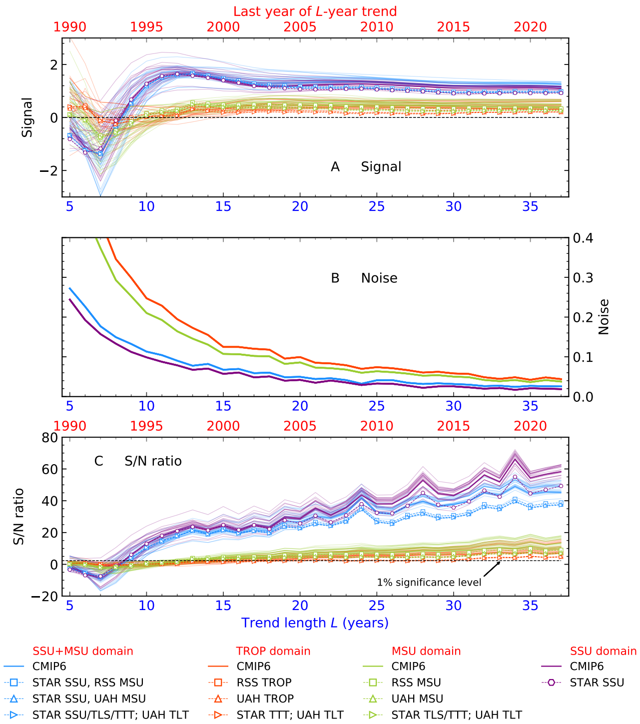

Fig. 5. Signal, noise, and S/N ratios in model and observed SSU and MSU data. Signals were calculated by projecting temperature data for different sets of atmospheric layers onto four fingerprints (SSU+MSU, TROP, MSU, and SSU) estimated from CMIP6 HIST$_{ext}$ simulations, and then fitting trends of increasing length L years to the resulting projection time series (panel A). CMIP6 control run temperature data were projected onto the same four fingerprints, yielding the projection time series $N_{ctl}(t)$. The noise $\sigma_{ctl}(L)$ is estimated by fitting non-overlapping L-year trends to $N_{ctl}(t)$ and calculating the standard deviation of the L-year trend distribution (panel B). The S/N ratio is the L-year signal in panel A divided by the respective values of $\sigma_{ctl}(L)$ in panel B (see Methods and SI). Model signals are from 32 HIST$_{ext}$ realizations; model noise is from 4,050 years of control run data. Signals and S/N ratios in which observed temperature data are used are plotted with symbols and dashed lines. The dashed horizontal line in panel C is the 1% significance level.

the spatial similarity between key features of the second EOF of the satellite data and certain CMIP6 HIST$_{ext}$ simulations, particularly the common negative loadings in the stratosphere at high latitudes of the Southern Hemisphere (see SI Fig. S4).

Sensitivity tests

We performed three sensitivity tests. The first explores the impact on fingerprint results of removing global-mean temperature signals. The second test considers the effect of accounting for large differences in the mass of the six atmospheric layers analyzed here. The third test examines whether S/N results are biased by overlap between the weighting functions used to sample the temperatures of these six layers (59).

In the first test, we find that removal of overall global-mean stratospheric cooling and tropospheric signals does not negate confident identification of an anthropogenic fingerprint in the vertical structure of atmospheric temperature change (SI Figs. S5 and S6). However, removing global-mean temperature changes in each of the six individual atmospheric layers – thereby removing information about vertical temperature-change gradients – markedly reduces S/N ratios and fingerprint detectability (see Methods and SI).

The second and third sensitivity tests are described in the Methods and SI. Although both tests reduce S/N values (see SI Figs. S7 and S8), the model-predicted SSU+MSU fingerprint can still be consistently identified in each of the individual HIST$_{ext}$ realizations and satellite data sets.

Conclusions

Our results illustrate that including information from the mid- to upper stratosphere (S_{25-50}) substantially enhances the detectability of an anthropogenic fingerprint on Earth's atmospheric temperature. This enhancement holds for observations and for individual model HIST$_{ext}$ realizations. Extending latitude-height fingerprints from the lower stratosphere to the S_{25-50} layer samples a region of the atmosphere where the direct radiative signature of CO_2 is prominent (1, 2, 8), the temperature signal driven by CO_2 increase is large, and the noise of natural internal variability is low.

The SSU+MSU vertical fingerprint extends from the lower troposphere to roughly 50 km above the surface. Signal-to-noise (S/N) ratios for the SSU+MSU domain consistently exceed 38 in the satellite data analyzed here. This value is virtually impossible to obtain by chance alone if our model-based estimates of signal and noise are realistic (55). In the satellite data sets, the S/N ratios for the SSU+MSU domain are roughly a factor of five larger than in the case of the "MSU only" vertical fingerprint, which truncates at an altitude of approximately 20-25 km (Fig. 5C).

The larger S/N values for the SSU+MSU fingerprint arise not only from the large cooling signal in the mid- to upper stratosphere, but also from the low internal variability noise in the S_{25-50} layer (Fig. 2) and the distinct differences between S_{25-50} signal and noise spatial patterns (SI Fig. S2). As a result, including the S_{25-50} layer in the SSU+MSU vertical fingerprint more effectively damps the noise of natural internal variability. A mass-weighted fingerprint analysis diminishes the contribution of stratospheric cooling and is less effective at separating signal and noise, but does not negate identification of the SSU+MSU fingerprint.

One issue revealed by this study warrants further attention. In the CMIP6 models analyzed here, model-predicted stratospheric cooling over 1986 to 2022 is significantly larger than in the SSU data (Figs. 2A-C). Multiple factors are likely to contribute to this discrepancy. These factors include model errors in the imposed anthropogenic and natural external forcings (42, 43, 60), in the simulated response to these forcings, and in the properties of internal variability. Mismatches in the random phasing of simulated and observed variability may also be relevant (41, 44), along with residual errors in satellite temperature data sets (25, 27, 61).

In the troposphere, accounting for model-versus-observed differences in the phasing of Pacific decadal variability improves agreement between simulated and observed temperature trends over the satellite era (44). The magnitude of decadal internal variability is smaller in the mid- to upper stratosphere than in the troposphere (Fig. 1). It is unlikely, therefore, that either phasing differences or model errors in the amplitude of decadal variability could fully explain why the simulated cooling of the S_{25-50} layer is significantly larger than observed (Figs. 2A-C). Forcing errors appear to be a more plausible explanation for this discrepancy, particularly in view of the substantial (and ongoing) evolution of forcing

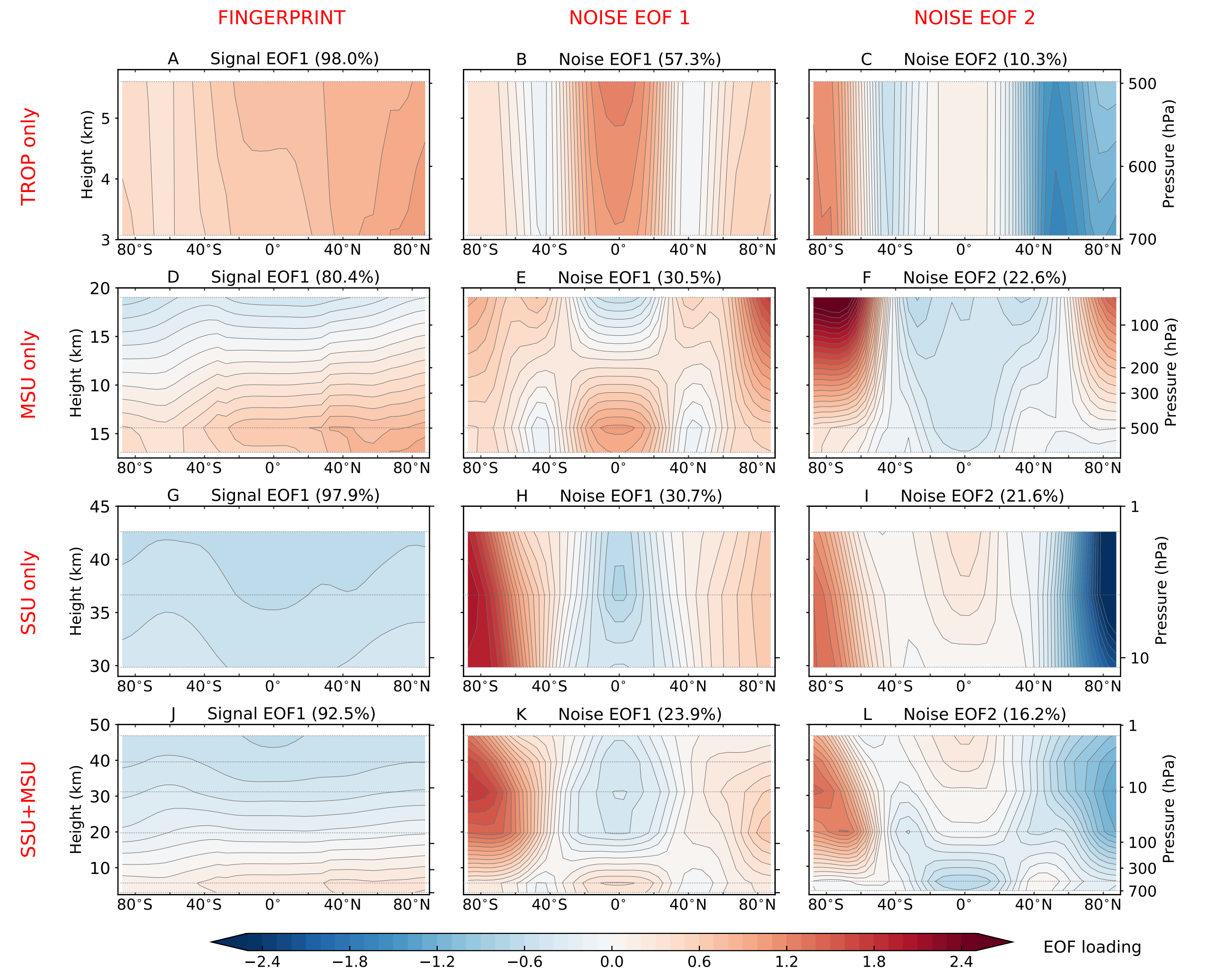

Fig. 6. Fingerprints and leading noise modes in CMIP6 simulations. Results are for four different spatial domains: TROP, MSU, SSU, and SSU+MSU (rows 1-4, respectively). These domains comprise the two tropospheric layers, the three MSU layers, the three SSU layers, and the six MSU+SSU layers. The fingerprint (left column) is the first Empirical Orthogonal Function (EOF) of the multi-model average atmospheric temperature changes computed from 32 HIST$_{ext}$ realizations performed with nine different CMIP6 models. The first two noise EOFs (middle and right columns) were calculated from 4,050 years of concatenated pre-industrial control run data generated with the same nine models. In estimating fingerprints and noise EOFs, global-mean temperature changes were retained for each of the six atmospheric layers considered. The dotted horizontal grey lines are plotted at the approximate peaks of the three SSU and three MSU weighting functions. See SI for further details.

estimates between CMIP5 and CMIP6 (39, 42, 60).

The challenge in interpreting differences between simulated and observed temperature trends lies in reliably quantifying the relative contributions of the multiple factors mentioned above. Such work will benefit from systematic exploration of uncertainties in radiative forcing (42, 60, 62, 63). It is also important to perform rigorous model-data comparisons of decadal variability for stratospheric temperature, as has been done for tropospheric temperature (55, 64).[§]

Model-based decadal variability estimates are an integral part of our fingerprint study. The reliability of these estimates underpins the credibility of our S/N ratios (Fig. 5). We note, however, that the CMIP6 models analyzed here would have to underestimate the observed (but uncertain) natural internal variability of stratospheric temperature by more than an order of magnitude in order to negate identification of an anthropogenic fingerprint in the SSU and SSU+MSU domains. We find no evidence that such an error exists (see Fig. 1).

In summary, the warming of the troposphere and cooling of the stratosphere across all latitudes is a unique fingerprint of greenhouse gas forcing. If tropospheric warming were solely due to solar activity, warming rather than cooling of the upper stratosphere would be expected (15, 23, 65). Alternatively, if stratospheric cooling and tropospheric warming at all latitudes – sustained over decades – were caused by internal variability alone, then similar patterns should sometimes emerge in the many long control runs of global models. This is not the case. Thus the ability to examine the vertical structure of atmospheric temperature changes is a powerful tool for separating human and natural effects on climate. Extending the reach of "vertical fingerprinting" from the lower troposphere to the upper stratosphere provides incontrovertible evidence of anthropogenic impact on Earth's climate.

[§] Such comparisons are hampered by the relatively short length of the observations and by the availability of only a single manifestation of forced and unforced temperature changes.

Materials and Methods

Satellite data. We rely on satellite data from three groups: Remote Sensing Systems (RSS) (66), the Center for Satellite Applications and Research (STAR) (61, 67), and the University of Alabama at Huntsville (UAH) (26). STAR is the only current source of spatially resolved temperature data for SSU channels 1, 2, and 3 (27). STAR, RSS, and UAH each supply MSU-based measurements of the temperatures of the lower stratosphere (TLS) and the mid- to upper troposphere (TMT). We apply a standard regression-based method to adjust TMT for the influence it receives from lower stratospheric cooling (68, 69), thereby obtaining the temperature of the total troposphere (TTT; see SI). Only RSS and UAH provide MSU estimates of the temperature of the lower troposphere (TLT). We "pair" STAR SSU data with UAH and RSS MSU data to generate two observed data sets spanning the lower troposphere to the upper stratosphere. Pairing STAR SSU, TLS, and TTT data with UAH TLT data yields a third observed data set (see SI).

Model data. The model synthetic SSU and MSU temperatures analyzed here are from phase 6 of the Coupled Model Intercomparison Project (CMIP6) (32). "Synthetic" indicates that the model results were calculated with weighting functions that facilitate direct comparison between satellite and model temperature changes (see SI).

The synthetic SSU and MSU temperatures are from three different types of numerical experiment: 1) Simulations with estimated historical changes in natural and anthropopogenic external forcings, which typically commence from January 1850 and end in December 2014; 2) Scenario runs with post-2014 changes in anthropogenic external forcings that are specified according to a Shared Socioeconomic Pathway which reaches radiative forcing of 8.5 W/m^2 by 2100 (SSP5-8.5); and 3) Pre-industrial control integrations with no year-to-year changes in external forcings.

The CMIP6 historical and scenario simulations consider not only the effects of CO_2 increases, but also include the radiative effects of changes in other greenhouse gases (70), anthropogenic aerosols, and solar and volcanic forcing. Temperatures from historical simulations and corresponding scenario runs were spliced together to permit comparison of model and observational results over 1986 to 2022. We refer to these as extended historical runs (HIST_{ext}; see SI). The CMIP6 model historical and SSP5-8.5 simulations used in our study are identified in Table S1. The control runs required for noise estimation are listed in Table S2. We analyzed a total of 32 HIST_{ext} realizations performed with nine different models and control runs generated with the same nine models.

Fingerprint and signal trends. We project zonal-mean annual-mean atmospheric temperature onto a searched-for fingerprint pattern $F(x,p)$ estimated from the multi-model average temperature changes in the HIST_{ext} simulations. This yields the projection time series $Z(t)$, a measure of uncentered spatial covariance (see SI). The indices x, p, and t are over latitude, atmospheric layer, and time (respectively). The $T(x,p,t)$ temperature data projected onto $F(x,p)$ are either from satellite observations or individual HIST_{ext} realizations. $Z(t)$ is a measure of the evolving pattern similarity between $F(x,p)$ and $T(x,p,t)$ at each year t. We compute L-year least-squares linear trends in $Z(t)$, starting in 1986, the beginning of continuous SSU records. The first trend length L is five years, corresponding to the period 1986 to 1990; L is increased in one-year increments, with $L = 37$ corresponding to 1986 to 2022. The signal $S(L)$ is the least-squares trend in $Z(t)$. Large $S(L)$ trends denote time-increasing similarity between the latitude-height temperature changes in $T(x,p,t)$ and the fingerprint pattern.

Noise trends. To determine whether and when the values of $S(L)$ in Fig. 5A achieve statistical significance, we compare $S(L)$ with null distributions in which we know *a priori* that natural internal variability is the only explanation for trends in pattern similarity. We use control runs with no year-to-year changes in external forcing to generate these "no signal" distributions. We project a total of 4,050 years of atmospheric temperature data from nine CMIP6 pre-industrial control runs onto the TROP, SSU, MSU, and SSU+MSU fingerprints, resulting in a projection time series $N_{ctl}(t)$ for each fingerprint. Non-overlapping L-year trends in $N_{ctl}(t)$ are then calculated for each value of L considered (i.e., for $L = 5, 6, \ldots 37$ years). For the $L = 37$-year analysis period, there are 109 individual samples of trends in $N_{ctl}(t)$. The standard deviation of these L-year noise trend distributions, $\sigma_{ctl}(L)$, is shown in Fig. 5B and is the denominator of the S/N ratios in Fig. 5C.

Global-mean removal. To determine whether our S/N results are solely driven by large global-mean temperature changes (21, 39), we compared the baseline case in Fig. 5 (Case 1, which includes global-mean changes) with two additional cases. In Case 2, the global-mean temperature change in each of the six layers was removed from each latitude band of each layer. Removal is performed for each year t and each model and observational data set. Case 3 is analogous to Case 2, but the subtraction involved the overall global-mean stratospheric temperature change (the average of the global-mean changes in the three SSU channels and TLS) and the overall global-mean tropospheric temperature change (the average of the global-mean changes in TTT and TLT). These sensitivity tests are described in the SI and are shown in SI Figs. S5 and S6 for the six-layer SSU+MSU domain.

ACKNOWLEDGMENTS. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison (PCMDI) provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. B.D.S was supported by the Francis E. Fowler IV Center for Ocean and Climate at Woods Hole Oceanographic Institution (WHOI). Research at Lawrence Livermore National Laboratory (LLNL) was performed under the auspices of U.S. Department of Energy Contract DE-AC52-07NA27344. S.P. and K.E.T. were supported through the PCMDI Project, which is funded by the Regional and Global Model Analysis Program of the Office of Science at the US Department of Energy. C.-Z.Z. was supported by the NOAA Joint Polar Satellite System (JPSS) Proving Ground and Risk Reduction (PGRR) Program under NOAA grant NA19NES4320002 (Cooperative Institute for Satellite Earth System Studies-CISESS) at the University of Maryland/ESSIC. C.-Z.Z. was also funded by the National Centers for Environmental Information (NCEI) Climate Data Record (CDR) Program. Q.F. was supported by NSF Grant AGS-2202812. S.S. was partly funded by NSF AGS grant 1848863. D.W.J.T. was supported by the NSF Climate and Large-Scale Dynamics division. For C.-Z.Z., the views, opinions, and findings contained in this paper are those of the authors and should not be construed as an official NOAA or U.S. Government position, policy, or decision. Jia-Rui Shi (WHOI) provided helpful comments on the manuscript.

1. S Manabe, RT Wetherald, Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci.* **24**, 241–259 (1967).
2. HF Goessling, S Bathiany, Why CO_2 cools the middle atmosphere – a consolidating model perspective. *Earth Syst. Dynam.* **7**, 697–715 (2016).
3. SB Fels, JD Mahlman, MD Schwarzkopf, RW Sinclair, Stratospheric sensitivity to perturbations in ozone and carbon dioxide: Radiative and dynamical response. *J. Atmos. Sci.* **37**, 2265–2297 (1980).
4. DJ Karoly, et al., An example of fingerprint detection of greenhouse climate change. *Cli. Dyn.* **10**, 97–105 (1994).
5. U Langematz, An estimate of the impact of observed ozone losses on stratospheric temperature. *Geophys. Res. Lett.* **27**, 2077–2080 (2000).
6. SM Rosier, KP Shine, The effect of two decades of ozone change on stratospheric temperature as indicated by a general circulation model. *Geophys. Res. Lett.* **27**, 2617–2620 (2000).
7. V Ramaswamy, MD Schwarzkopf, Effects of ozone and well-mixed gases on annual-mean stratospheric temperature trends. *Geophys. Res. Lett.* **29**, 21–1–21–4 (2002).
8. DWJ Thompson, et al., The mystery of recent stratospheric temperature trends. *Nature* **491**, 692–697 (2012).
9. AH Oort, H Liu, Upper-air temperature trends over the globe, 1958-1989. *J. Clim.* **6**, 292–307 (1993).
10. DE Parker, et al., A new global gridded radiosone temperature data base and recent temperature trends. *Geophys. Res. Lett.* **24**, 1499–1502 (1997).
11. RW Spencer, JR Christy, Precision and radiosonde validation of satellite gridpoint temperature anomalies. Part II: A tropospheric retrieval and trends during 1979-1990. *J. Clim.* **5**, 858–866 (1992).
12. FJ Wentz, M Schabel, Effects of orbital decay on satellite-derived lower-tropospheric temperature trends. *Nature* **394**, 661–664 (1998).
13. CZ Zou, et al., Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasses. *J. Geophys. Res.* **111**, D19114 (2006).

14. E Kalnay, et al., The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–471 (1996).
15. BD Santer, et al., A search for human influences on the thermal structure of the atmosphere. *Nature* **382**, 39–46 (1996).
16. KY Vinnikov, A Robock, RJ Stouffer, S Manabe, Vertical patterns of free and forced climate variations. *Geophys. Res. Lett.* **23**, 1801–1804 (1996).
17. SFB Tett, JFB Mitchell, DE Parker, MR Allen, Human influence on the atmospheric vertical temperature structure: Detection and observations. *Science* **274**, 1170–1173 (1996).
18. PW Thorne, et al., Assessing the robustness of zonal mean climate change detection. *Geophys. Res. Lett.* **29** (2002).
19. FC Lott, et al., Models versus radiosondes in the free atmosphere: A new detection and attribution analysis of temperature. *J. Geophys. Res. Atmos.* **118**, 2609–2619 (2013).
20. V Ramaswamy, et al., Anthropogenic and natural influences in the evolution of lower stratospheric cooling. *Science* **311**, 1138–1141 (2006).
21. BD Santer, et al., Influence of satellite data uncertainties on the detection of externally forced climate change. *Science* **300**, 1280–1284 (2003).
22. BD Santer, et al., Human and natural influences on the changing thermal structure of the atmosphere. *Proc. Nat. Acad. Sci.* **110**, 17235–17240 (2013).
23. BD Santer, et al., Identifying human influences on atmospheric temperature. *Proc. Nat. Acad. Sci.* **110**, 26–33 (2013).
24. C Mears, FJ Wentz, P Thorne, D Bernie, Assessing uncertainty in estimates of atmospheric temperature changes from MSU and AMSU using a Monte-Carlo technique. *J. Geophys. Res.* **116**, D08112 (2011).
25. C Mears, FJ Wentz, Sensitivity of satellite-derived tropospheric temperature trends to the diurnal cycle adjustment. *J. Clim.* **29**, 3629–3646 (2016).
26. RW Spencer, JR Christy, WD Braswell, UAH version 6 global satellite temperature products: Methodology and results. *Asia-Pac. J. Atmos. Sci.* **53**, 121–130 (2017).
27. CZ Zou, H Qian, Stratospheric temperature climate record from merged SSU and AMSU-A observations. *J. Atmos. Ocean. Tech.* **33**, 1967–1984 (2016).
28. S Solomon, et al., Mirrored changes in Antarctic ozone and stratospheric temperature in the late 20th versus early 21st centuries. *J. Geophys. Res.* **122**, 8940–8950 (2017).
29. CZ Zou, H Qian, W Wang, L Wang, C Long, Recalibration and merging of SSU observations for stratospheric temperature trend studies. *J. Geophys. Res.* **119**, 13180–13205 (2014).
30. J Nash, R Saunders, A review of stratospheric sounding unit radiance observations for climate trends and reanalyses. *Q. J. Roy. Met. Soc.* **141**, 2103–2113 (2015).
31. AC Maycock, et al., Revisiting the mystery of recent stratospheric temperature trends. *Geophys. Res. Lett.* **45**, 9919–9933 (2018).
32. V Eyring, et al., Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Mod. Dev.* **9(5)**, 1937–1958 (2016).
33. K Hasselmann, *On the signal-to-noise problem in atmospheric response studies.* (Roy. Met. Soc., London), pp. 251–259 (1979).
34. GC Hegerl, et al., Detecting anthropogenic climate change with an optimal fingerprint method. *J. Clim.* **9**, 2281–2306 (1996).
35. MC Casas, et al., Understanding model-observation discrepancies in satellite retrievals of atmospheric temperature using GISS ModelE. *J. Geophys. Res.* **128**, e2022JD037523 (2022).
36. J Rao, CI Garfinkel, CMIP5/6 models project little change in the statistical characteristics of sudden stratospheric warmings in the 21st century. *Environ. Res. Lett.* **16**, 034024 (2021).
37. BD Santer, et al., Signal-to-noise analysis of time-dependent greenhouse warming experiments. *Cli. Dyn.* **9**, 267–285 (1994).
38. DM Mitchell, YTE Lo, WJM Seviour, L Haimberger, LM Polvani, The vertical profile of recent tropical temperature trends: Persistent model biases in the context of internal variability. *Env. Res. Lett.* **15**, 1040b4 (2020).
39. BD Santer, et al., Using climate model simulations to constrain observations. *J. Clim.* **34**, 6281–6301 (2021).
40. BD Santer, et al., Volcanic contribution to decadal changes in tropospheric temperature. *Nat. Geosci.* **7**, 185–189 (2014).
41. S Po-Chedley, et al., Natural variability drives model-observational differences in tropical tropospheric warming. *Proc. Nat. Acad. Sci.* **118**, e2020962118 (2021).
42. JC Fyfe, V Kharin, BD Santer, RNS Cole, NP Gillett, Significant impact of forcing uncertainty in a large ensemble of climate model simulations. *Proc. Nat. Acad. Sci.* **118**, e2016549118 (2021).
43. JT Fasullo, et al., Spurious late historical-era warming in CESM2 driven by prescribed biomass burning emissions. *Geophys. Res. Lett.* **49**, e2021GL097420 (2022).
44. S Po-Chedley, et al., Internal variability and forcing influence model-satellite differences in the rate of tropical tropospheric warming. *Proc. Nat. Acad. Sci.* **119**, e2209431119 (2022).
45. S Solomon, et al., Emergence of healing in the Antarctic ozone layer. *Science* **353**, 269–274 (2016).
46. V Ramaswamy, MD Schwarzkopf, WJ Randel, Fingerprint of ozone depletion in the spatial and temporal pattern of recent lower-stratospheric cooling. *Nature* **382**, 616–618 (1996).
47. P Forster, et al., Changes in Atmospheric Constituents and in Radiative Forcing in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, eds. S Solomon, et al. (Cambridge University Press), pp. 129–234 (2007).
48. G Myhre, et al., Anthropogenic and natural radiative forcing in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds. TF Stocker, et al. (Cambridge University Press), pp. 659–740 (2013).
49. C Bonfils, et al., Human influence on joint changes in temperature, rainfall and continental aridity. *Nat. Clim. Chang.* **10**, 726–731 (2020).
50. K Riahi, et al., The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Glob. Env. Chang.* **42**, 153–168 (2017).
51. Q Fu, P Lin, S Solomon, DL Hartmann, Observational evidence of strengthening of the Brewer-Dobson circulation since 1980. *J. Geophys. Res.* **120**, 10214–10228 (2015).
52. N Feldl, S Po-Chedley, HKA Singh, S Hay, PJ Kushner, Sea ice and atmospheric circulation shape the high-latitude lapse rate feedback. *npj. Clim. Atmos. Sci.* **41** (2020).
53. M Rantanen, et al., The Arctic has warmed nearly four times faster than the globe since 1979. *Commun. Earth. Environ.* **3** (2022).
54. J Marshall, et al., The ocean's role in polar climate change: asymmetric Arctic and Antarctic responses to greenhouse gas and ozone forcing. *Phil. Trans. Roy. Soc. A* **372** (2014).
55. BD Santer, et al., Robust anthropogenic signal identified in the seasonal cycle of tropospheric temperature. *J. Clim.* **35**, 6075–6100 (2022).
56. BD Santer, et al., Separating signal and noise in atmospheric temperature changes: The importance of timescale. *J. Geophys. Res.* **116**, D22105 (2011).
57. E Hawkins, R Sutton, The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Met. Soc.* **90**, 1095–1108 (2009).
58. WJ Randel, AK Smith, F Wu, CZ Zou, H Qian, Stratospheric temperature trends over 1979–2015 derived from combined SSU, MLS, and SABER satellite observations. *J. Clim.* **29**, 4843–4859 (2016).
59. A Steiner, et al., Observed temperature changes in the troposphere and stratosphere from 1979 to 2018. *J. Clim.* **33**, 8165–8194 (2020).
60. LA Rieger, WJ Randel, AE Bourassa, S Solomon, Stratospheric temperature and ozone anomalies associated with the 2020 Australian New Year fires. *Geophys. Res. Lett.* **48**, e2021GL095898 (2021).
61. CZ Zou, H Xu, X Hao, Q Liu, Mid-tropospheric layer temperature record derived from satellite microwave sounder observations with backward merging approach. *J. Geophys. Res.* **128**, e2022JD037472 (2023).
62. R Pincus, PM Forster, B Stevens, The Radiative Forcing Model Intercomparison Project (RFMIP): Experimental protocol for CMIP6. *Geosci. Mod. Dev.* **9**, 3447–3460 (2017).
63. NP Gillett, et al., The detection and attribution model intercomparison project (DAMIP v1.0) contribution to CMIP6. *Geosci. Mod. Dev.* **9**, 3685–3697 (2016).
64. J Pallotta, BD Santer, Multi-frequency analysis of simulated versus observed variability in tropospheric temperature. *J. Clim.* **33**, 10383–10402 (2020).
65. GC Hegerl, et al., Understanding and Attributing Climate Change in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, eds. S Solomon, et al. (Cambridge University Press), pp. 663–745 (2007).
66. C Mears, FJ Wentz, A satellite-derived lower-tropospheric atmospheric temperature dataset using an optimized adjustment for diurnal effects. *J. Clim.* **30**, 7695–7718 (2017).
67. CZ Zou, MD Goldberg, X Hao, New generation of U.S. satellite microwave sounder achieves high radiometric stability performance for reliable climate change detection. *Sci. Adv.* **4**, eaau0049 (2018).
68. Q Fu, CM Johanson, SG Warren, DJ Seidel, Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature* **429**, 55–58 (2004).
69. Q Fu, CM Johanson, Stratospheric influences on MSU-derived tropospheric temperature trends: A direct error analysis. *J. Clim.* **17**, 4636–4640 (2004).
70. B Govindasamy, et al., Limitations of the equivalent CO_2 approximation in climate change simulations. *J. Geophys. Res.* **106**, 22593–22603 (2001).

Supporting Information for

Exceptional Stratospheric Contribution to Human Fingerprints on Atmospheric Temperature

Benjamin D. Santer et al.

Corresponding Author: Benjamin D. Santer.
E-mail: bensanter1289@gmail.com

This PDF file includes:

Supporting text
Figs. S1 to S8
Tables S1 to S2
SI References

Supporting Information Text

Materials

Additional information on satellite data. We rely on estimates of the temperature of the lower stratosphere (TLS), mid-troposphere (TMT), and lower troposphere (TLT) derived from satellite-borne Microwave Sounding Units (MSU) and Advanced Microwave Sounding Units (AMSU). These data sets are produced by Remote Sensing Systems (RSS) (1) and the University of Alabama at Huntsville (UAH) (2). We also use TLS and TMT data from the Center for Satellite Applications and Research (STAR) (3, 4). STAR does not currently provide TLT data.

Information on temperature changes in the mid- to upper stratosphere is available from channels 1, 2, and 3 of the Stratospheric Sounding Unit (SSU). The SSU temperature data are from STAR (5). We use the most recent versions of the MSU/AMSU and SSU/AMSU-A data:

- RSS 4.0 and UAH 6.0 for TLS, TMT, and TLT;
- STAR 5.0 for TLS and TMT;
- STAR 3.0 for SSU1, SSU2, and SSU3.

Version 3 of the STAR SSU data merged the version 2 SSU data set (6) with 8 channels of AMSU-A observations. Merging extends the SSU time series from 2006 to present (5). MSU data are merged with AMSU data after 1998. We refer to these merged products subsequently as "SSU" and "MSU".

We employed a standard regression-based method to adjust TMT for the influence it receives from lower stratospheric cooling (7). This adjustment yields TTT, the temperature of the "total" troposphere (see SI section "Method for correcting TMT data").

Our fingerprint analysis employs zonally averaged temperature changes for SSU3, SSU2, SSU1, TLS, TTT, and TLT. The approximate peaks of the weighting functions for these layers are 45, 38, 30, 19, 5.6, and 3.1 km, respectively.

All satellite temperature data sets analyzed here are in the form of monthly means on the same $2.5^\circ \times 2.5^\circ$ latitude/longitude grid. At the time this analysis was performed, satellite temperature data for full 12-month years were available for the 528-month period from January 1979 to December 2022 for TLS, TTT, and TLT and for the 444-month period from January 1986 to December 2022 for SSU3, SSU2, and SSU1. We use the latter period here since we require non-missing temperature data over a common time window for all six layers of interest.

As noted above, STAR does not have a TLT product. To include STAR MSU data in our study, we first calculated TTT from STAR TLS and TMT data, and then generated data sets in which the STAR SSU, TLS and TTT data were "paired" with either RSS TLT or UAH TLT:

$$\begin{aligned} \text{STAR1} &= \text{STAR SSU3/2/1} + \text{STAR TLS/TTT} + \text{RSS TLT} \\ \text{STAR2} &= \text{STAR SSU3/2/1} + \text{STAR TLS/TTT} + \text{UAH TLT} \end{aligned}$$

Relative to STAR1, S/N ratios obtained with STAR2 data are approximately 30% smaller for the TROP case (because the lower tropospheric warming is smaller in UAH than in RSS; see main text Fig. 2). This means that for the TROP domain, S/N ratios estimated with STAR2 data are more conservative. Nevertheless, the model-predicted TROP fingerprints can be identified at the 1% level in both the STAR1 and STAR2 observational temperature data sets.

Whether we use STAR1 or STAR2 has minimal impact on S/N results for the SSU+MSU and MSU domains. This lack of sensitivity is due the fact that the TLT layer is only one-sixth and one-third of the SSU+MSU and MSU domains (respectively). In the main text (in Fig. 5) and in Figs. S2, S5, S7, and S8) we show STAR2 results only.

Additional information on model data. We analyze synthetic SSU3, SSU2, SSU1, TLS, TTT, and TLT data from simulations performed under phase 6 of the Coupled Model Intercomparison Project (CMIP6) (8). "Synthetic" denotes the calculation of a vertically weighted average of atmospheric temperature in order to facilitate the comparison of simulations and satellite SSU or MSU data (see SI section "Calculation of synthetic satellite temperatures"). The synthetic SSU and MSU temperatures are from three different types of numerical experiment:

1. Simulations with estimated historical changes in natural and anthropopogenic external forcings, which typically commence from January 1850 and end in December 2014.
2. Scenario runs with post-2014 changes in anthropogenic external forcings that are specified according to a Shared Socioeconomic Pathway (SSP). The SSP used here is referred to as SSP5-8.5 (or as SSP5) because it reaches radiative forcing of 8.5 W/m^2 by 2100. We adopt the SSP5 nomenclature here (9).
3. Preindustrial control integrations with no year-to-year changes in external forcings.

Each historical simulation was spliced together with a companion SSP5 run initiated from the end of the historical run. This extension of the historical run allows us to compare simulated and observed atmospheric temperatures over the full period with continuous availability of monthly-mean MSU and SSU data (1986 to 2022; see SI section "Additional information on satellite data"). We refer to these subsequently as HIST$_{ext}$ runs.

To calculate synthetic SSU data, we require simulation output from CMIP6 models with sufficient vertical resolution in the mid- to upper stratosphere. We follow the recommendations of Thompson et al. here (10) and require models with a top located at 0.1 hPa or higher in order to compute synthetic temperatures for all three SSU channels. Output fulfilling this requirement is available from models participating in the Aerosols and Chemistry Model Intercomparison Project (AerChemMIP) (11). Here, we use the AerChemMIP "plev39" data with zonal-mean monthly-mean atmospheric temperatures at 39 standard pressure levels.*

In addition to the requirement of a sufficiently high top, there were three further requirements for inclusion of a CMIP6 model in the fingerprint analysis. First, given the large warming signatures of major volcanic eruptions on stratospheric temperatures (10, 12), only models that explicitly included the full radiative effects of volcanic aerosols were considered (13). Neglecting the large effect of the 1991 Pinatubo eruption would bias comparisons between simulated and observed stratospheric temperature changes over 1986 to 2022. Second, any model with spurious variability in stratospheric temperature was excluded.†

Finally, we required that the data for computing synthetic MSU temperatures had to exist for the same simulations from which we had calculated synthetic SSU temperatures. These three requirements were satisfied in 32 different HIST_{ext} realizations performed with 9 different CMIP6 models. We analyzed control integrations from the same 9 models. Details of the model HIST_{ext} and control simulations are given in Tables S1 and S2, respectively.

Methods

Calculation of synthetic satellite temperatures. We used a local weighting function method developed at RSS to calculate synthetic MSU temperatures from the CMIP6 HIST_{ext} and preindustrial control runs (15). At each grid-point, simulated temperature profiles were convolved with local weighting functions. Weights depend on the grid-point surface pressure, the surface type (land, ocean, or sea ice), and the selected satellite channel (TLS, TMT, or TLT).

Because the influence of topography on weighting functions is not important in the mid- to upper stratosphere, use of a local weighting function method is not necessary for calculating synthetic SSU temperatures. We applied weighting functions available from STAR (5) to the zonal-mean monthly-mean plev39 atmospheric temperature data (see SI section "Additional information on model data") in order to derive synthetic SSU1, SSU2, and SSU3 data.

Method for correcting TMT data. Trends in TMT estimated from microwave sounders receive a substantial contribution from the cooling of the lower stratosphere (7). This contribution hampers reliable interpretation of the warming of the free troposphere – which is why most analysts adjust satellite TMT measurements and model simulations of TMT for the influence of stratospheric cooling (14–21).

An additional complication in comparing and interpreting uncorrected TMT results is that stratospheric cooling can vary appreciably in different observational data sets (22) and in different climate models (14, 15). In models, this is often due to large differences in stratospheric ozone forcing over the satellite era (13), or to systematic changes in stratospheric ozone forcing between different generations of CMIP models (14, 23).

Adjustment of TMT using the regression-based method introduced by Fu et al. (7) simplifies the interpretation of data-data, model-model, and model-data comparisons of tropospheric temperature change.‡ This method has been validated with both observed and model atmospheric temperature data (16, 24, 25).

In the following, we refer to adjusted TMT as total tropospheric temperature (TTT). It is calculated as follows:

$$\text{TTT} = a_{24}\text{TMT} + (1 - a_{24})\text{TLS} \qquad [1]$$

We compute two different versions of total tropospheric temperature: TTT_1 and TTT_2. TTT_1 was first used for adjusting tropical averages of TMT, with $a_{24} = 1.1$ at each latitude (17). In TTT_2, $a_{24} = 1.1$ between 30°N and 30°S, and $a_{24} = 1.2$ poleward of 30°.

The advantage of TTT_2 is that lower stratospheric cooling makes a larger contribution to unadjusted TMT trends at mid- to high latitudes. The latitudinally varying regression coefficients in TTT_2 remove more of this extratropical cooling. We use TTT_2 throughout the main text and the SI, and do not use the subscript "2" to identify TTT_2.

In practice, whether we use TTT_1 or TTT_2 has minimal influence on our S/N results.

We note that TTT_2 is calculated in the same way in all simulations and observations and for all months. This ensures that model-versus-observed temperature comparisons of TTT_2 are not affected by the application of regression coefficients that differ in the CMIP6 simulations and in satellite data.

Fingerprint analysis. Detection methods generally require an estimate of the true but unknown climate-change signal in response to an individual forcing or set of forcings (26). This is often referred to as the fingerprint, which we denote here by $F(x, p)$, where x is an index over latitude and p is an index over atmospheric layers.

Fingerprints can be defined in different ways. Here, $F(x, p)$ is the first Empirical Orthogonal Function (EOF) of the multi-model ensemble-mean change in temperature across the CMIP6 HIST_{ext} simulations.

*The plev39 levels (in hPa) are 1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 170, 150, 130, 115, 100, 90, 80, 70, 50, 30, 20, 15, 10, 7, 5, 3, 2, 1.5, 1, 0.7, 0.5, 0.4, 0.3, 0.2, 0.15, 0.1, 0.07, 0.05, and 0.03. For further details, see https://cmip6dr.github.io/Data_Request_Home/Documents/CMIP6_pressure_levels.pdf

†This is the case with CanESM5, which "exhibits anomalous aperiodic 1–2-month lower-stratospheric warming events in certain ensemble members" (14).

‡For example, differences between simulated and observed trends in unadjusted TMT could arise from the combined effects of model climate sensitivity errors (which would affect tropospheric temperature) and from unrelated model errors in stratospheric ozone forcing (which would primarily affect lower stratospheric temperature). Use of adjusted TMT reduces the contribution of stratospheric ozone forcing errors to model-versus-data differences in tropospheric temperature trends.

Let $T_{hst}(i, j, x, p, t)$ represent the temperature anomaly for the i^{th} HIST$_{ext}$ realization of the j^{th} CMIP6 model, where:

$$
\begin{array}{lcll}
i & = & 1, \ldots N_r(j) & \text{(no. of HIST}_{ext}\text{ realizations for the } j^{th} \text{ model)} \\
j & = & 1, \ldots N_{mod} & \text{(no. of CMIP6 models used in the fingerprint analysis)} \\
x & = & 1, \ldots N_x & \text{(no. of latitude bands with zonal-mean temperatures)} \\
p & = & 1, \ldots N_p & \text{(total no. of SSU and MSU atmospheric layers)} \\
t & = & 1, \ldots N_t & \text{(time in years)}
\end{array}
$$

Here, $N_r(j)$ varies from 1 to 10 realizations and $N_{mod} = 9$. After transforming synthetic MSU temperature data from each model's native grid to a common 5°× 5° latitude/longitude grid and calculating zonal averages, $N_x = 36$ latitude bands. Synthetic SSU data (which are already in zonal-mean form; see SI section "Additional infoprmation on model data") are transformed to the same 36 latitude nodes. N_p varies from 2 to 6 layers (see below). Fingerprint estimation is over the period of common coverage in SSU and MSU (1986 to 2022), so N_t is 37 years.

Anomalies in $T_{hst}(i, j, x, p, t)$ were defined relative to climatological annual means over 1986 to 2022. The multi-model ensemble-mean change, $\overline{T_{hst}}(x, p, t)$, was calculated by first averaging over the $N_r(j)$ individual realizations in the j^{th} model and then averaging over all N_{mod} models. The fingerprint $F(x, p)$ is the first EOF of $\overline{T_{hst}}(x, p, t)$. The time period used for determining $T_{obs}(x, p, t)$, the change in zonal-mean annual-mean atmospheric temperature in a selected combination of observed SSU and MSU data sets, is the same as used for calculating the fingerprint (1986 to 2022).

We estimate one fingerprint for each of the four different sets of the six atmospheric layers considered here:

1. SSU+MSU (six layers; SSU3, SSU2, SSU1, TLS, TTT, and TLT);
2. TROP (two layers; TTT and TLT);
3. MSU (three layers; TLS, TTT, and TLT);
4. SSU (three layers; SSU3, SSU2, and SSU1).

The TROP and SSU cases provide information on the S/N properties of satellite era temperature changes in the troposphere and in the mid- to upper stratosphere (respectively). Comparison of S/N results for the MSU and SSU+MSU domains yields insights into the impact of extending previous "vertical fingerprint" studies to the upper stratosphere. Previous studies were conducted using MSU information only (27) and were therefore restricted to the troposphere and lower stratosphere.

For each of these four different sets of atmospheric layers, we seek to determine whether the pattern similarity between $F(x, p)$ and $T_{obs}(x, p, t)$ shows a statistically significant increase over time. We also consider whether there is a significant increase in pattern similarity between the fingerprint and each individual HIST$_{ext}$ realization – i.e., between $F(x, p)$ and $T_{hst}(i, j, x, p, t)$.

To address these two questions, we require control run estimates of internally generated variability in which we know *a priori* that there is no expression of the fingerprint, except by chance. We obtain such variability estimates from control runs performed with the same nine CMIP6 models used to estimate $F(x, p)$. Layer-average atmospheric temperatures from each control run are regridded to the same 5°× 5° latitude/longitude grid used for fingerprint estimation. After regridding and calculation of zonal averages, layer-average atmospheric temperature anomalies are defined relative to climatological annual means computed over the full length of each control run.

Because the length of the nine CMIP6 control runs varies by a factor of approximately 2 (see Table S2), models with longer control integrations could have a disproportionately large impact on our noise estimates. To guard against this possibility, we rely on the last 450 years of each model's pre-industrial control run. Use of the last 450 years reduces the contribution of initial residual drift and guarantees that each model is given equal weight in calculating the denominator of our S/N ratios. Concatenation yields 9 × 450 = 4,050 years of control run atmospheric temperature output.

Use of the last 450 years of each control run may not fully remove non-physical residual drift, which can inflate and bias S/N estimates (28). Here, we assume that drift behavior can be well-approximated by a least-squares linear trend and the drift is removed at each latitude band and for each atmospheric layer. Drift removal is performed over the last 450 control run years only (since only the last 450 years are concatenated).

In processing the observations, layer-average atmospheric temperature data from STAR, RSS, and UAH are first regridded to the same target 5°× 5° latitude/longitude grid used for the model HIST$_{ext}$ simulations and control runs. Observations are then zonally averaged and expressed as anomalies relative to climatological annual means over 1986 to 2022. The observed temperature anomaly data, $T_{obs}(x, p, t)$, are then projected onto $F(x, p)$, the time-invariant fingerprint:

$$Z_{obs}(t) = \sum_{x=1}^{N_x} \sum_{p=1}^{N_p} T_{obs}(x, p, t)\, F(x, p) \qquad [2]$$

$$t = 1, \ldots, 37.$$

This projection is equivalent to a spatially uncentered covariance between the $T_{obs}(x,p,t)$ and $F(x,p)$ patterns at year t. The signal time series $Z_{obs}(t)$ provides information on the fingerprint strength in the observations. If $T_{obs}(x,p,t)$ is becoming increasingly similar to $F(x,p)$, $Z_{obs}(t)$ should increase over time.

The projection of an individual HIST$_{ext}$ realization onto $F(x,p)$ is defined analogously:

$$Z_{hst}(i,j,t) = \sum_{x=1}^{N_x}\sum_{p=1}^{N_p} T_{hst}(i,j,x,p,t)\,F(x,p)$$
$$i = 1,\ldots,N_r(j);\;\; j = 1,\ldots,N_{mod}\,;\;\; t = 1,\ldots,37. \qquad [3]$$

To assess the significance of the changes in $Z_{obs}(t)$ or in $Z_{hst}(i,j,t)$, we compare trends in $Z_{obs}(t)$ and in $Z_{hst}(i,j,t)$ with a null distribution of trends. To generate a suitable null distribution, we require a case in which $T_{obs}(x,p,t)$ or $T_{hst}(i,j,x,p,t)$ is replaced by a record in which we know *a priori* that there is no expression of the fingerprint, except by chance. Here, we use a concatenated multi-model noise data set, $T_{ctl}(x,p,t)$, which has been regridded and detrended as described above.[§] The noise time series $N_{ctl}(t)$ is the projection of $T_{ctl}(x,p,t)$ onto the fingerprint:

$$N_{ctl}(t) = \sum_{x=1}^{N_x}\sum_{p=1}^{N_p} T_{ctl}(x,p,t)\,F(x,p)$$
$$t = 1,\ldots,N_{t\{ctl\}}. \qquad [4]$$

where $N_{t\{ctl\}}$ is 4,050, the total number of years in the multi-model noise estimate.

As in our previous work (29, 30), we fit least-squares linear trends of increasing length L years to $Z_{obs}(t)$. This yields $S_{obs}(L)$. We then form the signal-to-noise ratios SN$_{obs}(L)$ by dividing $S_{obs}(L)$ by $\sigma_{ctl}(L)$, the standard deviation of the distribution of non-overlapping L-length noise trends in $N_{ctl}(t)$. Signal trends in $Z_{hst}(i,j,t)$ are treated analogously – i.e., we calculate $S_{hst}(i,j,L)$ from $Z_{hst}(i,j,t)$, divide $S_{hst}(i,j,L)$ by $\sigma_{ctl}(L)$, and obtain SN$_{hst}(i,j,L)$.

We assess statistical significance by comparing these calculated S/N ratios with a Gaussian distribution, as in (31). This assumes that L-year trends in $N_{ctl}(t)$ have a Gaussian distribution. This assumption is reasonable for multi-model estimates of internal variability given the large sample sizes that we have here. Signal detection is stipulated to occur at the trend length L_d for which the S/N ratio first exceeds some stipulated significance level (typically 1% here) and then remains above that level for all values of $L > L_d$. The test is one-tailed.

Empirical estimates of the significance of our S/N ratios yield very similar results. These estimates are based on comparisons of signal trends with the actual distributions of L-year noise trends obtained from $N_{ctl}(t)$.

The start date for fitting linear trends to $Z_o(t)$ is 1986, the first complete year of common continuous temporal coverage of the observational SSU and MSU data. We use a minimum trend length of 5 years, so the first S/N ratio (and the earliest possible detection time) is for 5-year trends ending in 1990. The analysis period increases in increments of one year, i.e., $L = 5$, 6, 7, . . . , 37. The $L = 37$ case corresponds to the full satellite era (1986 to 2022).

Finally, we note that all model and observational temperature data used in the fingerprint analysis are appropriately area-weighted. Weighting involves multiplication by the square root of the cosine of the grid node's latitude (32). For visual display purposes only, the EOFs shown in Fig. 6 of the main text and in Figs. S3, S4, and S6 are unweighted (*i.e.*, the grid-point values of each EOF are divided by the square root of the cosine of the grid node's latitude). There is no weighting of the individual atmospheric layers – each layer has equal weight. Mass-weighted fingerprint results are discussed below (see SI section on "Mass and area weighting").

The S/N analysis described in the main text relies on the HIST$_{ext}$ fingerprints of zonal-mean annual-mean atmospheric temperature change. The CMIP6 HIST$_{ext}$ simulations involve combined anthropogenic and natural external forcing. Because anthropogenic forcing is substantially larger than natural external forcing over 1986 to 2022, the HIST$_{ext}$ fingerprints are very similar to fingerprint patterns obtained from integrations with anthropogenic forcing only (33). The HIST$_{ext}$ fingerprint patterns primarily reflect the tropospheric warming in response to human-caused changes in greenhouse gases and the stratospheric cooling caused by anthropogenic CO_2 increases and stratospheric ozone depletion (33).

For the SSU+MSU and SSU domains, the timescale-dependent S/N ratios in Fig. 5C of the main text show strong correlations across individual HIST$_{ext}$ realizations, despite the fact that the internal variability in each realization should not be correlated (except by chance). The explanation for this correlation across realizations is that the $S_{hst}(i,j,L)$ signals for the SSU+MSU and SSU domains are very large relative to the amplitude of the $\sigma_{ctl}(L)$ noise for these domains (compare Figs. 5A and B in the main text). This is why relatively small "noise" in the decay of $\sigma_{ctl}(L)$ as a function of increasing L, arising from our use of non-overlapping trends to estimate $\sigma_{ctl}(L)$, has large impact on SN(L) values and imparts correlation to SN(L) across the 32 HIST$_{ext}$ realizations.

[§] Unlike $T_{hst}(i,j,x,p,t)$, $T_{ctl}(x,p,t)$ has no index over i or over j. This is because there is typically only one realization of each control run and because the noise data from each of the 9 models have been concatenated.

Removal of spatial means. In comparing simulated and observed patterns of atmospheric temperature change and interpreting S/N results, we are interested in assessing contributions to S/N ratios from global- and from sub-global spatial scales. Our "baseline" fingerprint analysis in Fig. 5 of the main text relies on an uncentered spatial covariance statistic which retains the spatial means of the two fields that are being compared. The baseline case, therefore, incorporates both the global- and the sub-global components of temperature change.

As in our previous fingerprint work (34), it is of interest to determine whether large global-mean tropospheric warming and stratospheric cooling signals are the main driver of our consistent identification of model-predicted $F(x,p)$ fingerprints in satellite observations and in individual model HIST$_{ext}$ realizations (see Fig. 5C in main text). We address this question by comparing S/N ratios for the baseline case (Case 1, which includes global-mean temperature changes at each atmospheric level) with S/N results from two additional types of calculation:

1. For each of the N_p layers, N_x latitude bands, and N_t years, we remove the global-mean atmospheric temperature change for that layer, latitude band, and year (Case 2);

2. The overall global-mean tropospheric temperature change in year t (the average of the global-mean temperature changes for TTT and TLT in year t) is removed from the individual TTT and TLT layers. A similar subtraction is performed for each of the four stratospheric layers (SSU3, SSU2, SSU1, and TLS) using the overall global-mean stratospheric temperature change in year t (Case 3).

For example, for the observational zonal-mean annual-mean atmospheric temperature change used in Case 2:

$$<T_{obs}(p,t)> = \sum_{x=1}^{N_x} T_{obs}(x,p,t)\, W(x) \,/ \sum_{x=1}^{N_x} W(x) \qquad p = 1,\ldots,N_p; \quad t = 1,\ldots,37. \tag{5}$$

where $<T_{obs}(p,t)>$ is the global-mean temperature change for layer p and year t, the angle brackets denote a spatial average, and $W(x)$ are area weights for each latitude band. Subtraction of the global-mean temperature change yields:

$$T_{obs}(x,p,t)^{*} = T_{obs}(x,p,t) - <T_{obs}(p,t)> \qquad x = 1,\ldots,N_x; \quad p = 1,\ldots,N_p; \quad t = 1,\ldots,37. \tag{6}$$

where * denotes departures from the global-mean.

In Case 3, $< T_{obs}\{\textsc{strat}\}(t) >$ and $< T_{obs}\{\textsc{trop}\}(t) >$ are the overall global-mean temperature changes for the four stratospheric layers and the two tropospheric layers, respectively. These are removed from the individual stratospheric and tropospheric layers as follows:

$$\begin{aligned} T_{obs}(x,p,t)^{**} &= T_{obs}(x,p,t) - <T_{obs}\{\textsc{strat}\}(t)> \\ & x = 1,\ldots,N_x; \quad p = 1,\ldots,4; \quad t = 1,\ldots,37. \\ T_{obs}(x,p,t)^{**} &= T_{obs}(x,p,t) - <T_{obs}\{\textsc{trop}\}(t)> \\ & x = 1,\ldots,N_x; \quad p = 5,6; \quad t = 1,\ldots,37. \end{aligned} \tag{7}$$

where it is assumed that the ordering of layers is from the highest layer to the lowest layer and that the ordering of layers is identical in each data set, i.e., $p = 1$ is SSU3, $p = 2$ is SSU2, $p = 3$ is SSU1, $p = 4$ is TLS, $p = 5$ is TTT, and $p = 6$ is TLT. The double asterisk notation denotes a departure from the overall stratospheric or tropospheric global-mean (*c.f.* the single asterisk notation for Case 2).

While equations (5) though (7) are for observations, the processing is similar for HIST$_{ext}$ and for control simulations. In each model HIST$_{ext}$ or control run data set processed, we remove the global-mean temperature change for layer p from each latitude band of that layer (Case 2), or we remove the overall global-mean stratospheric temperature change from each latitude of each stratospheric layer and we subtract the overall global-mean tropospheric temperature change from each latitude of each tropospheric layer (Case 3).

For the HIST$_{ext}$ runs, these two different global-mean subtraction methods yield the multi-model ensemble means $\overline{T_{hst}}(x,p,t)^{*}$ (Case 2) and $\overline{T_{hst}}(x,p,t)^{**}$ (Case 3). The Case 2 fingerprint shown in Fig. S6B is $F(x,p)^{*}$, the leading EOF of $\overline{T_{hst}}(x,p,t)^{*}$. The Case 3 fingerprint in Fig. S6C is $F(x,p)^{**}$, the leading EOF of $\overline{T_{hst}}(x,p,t)^{**}$.

The key difference between Case 2 and Case 3 is that in the latter, we retain global-scale signals of interest in the observations and HIST$_{ext}$ runs, such as the increase in the size of stratospheric cooling with increasing altitude in the stratosphere (35) and the amplification of tropical tropospheric warming in TTT relative to TLT (20, 36). These global-scale signals are removed in Case 2.

Mass and area weighting. The focus of our study is on the value of including the mid- to upper stratosphere in climate fingerprinting. We seek to determine whether including temperature information from the S_{25-50} layer aids in separating anthropogenic climate change from natural internal variability. To address this question, each of the six atmospheric layers considered here was assigned a vertical weight of 1 in the fingerprint analysis. With uniform vertical weighting, including the S_{25-50} layer significantly enhances our ability to discriminate between human-caused climate change and internal variability (see Fig. 5C in the main text).

To explore the impact of mass weighting on our fingerprint results, we require a set of suitable weights that reflect the sampling of atmospheric mass by the weighting functions of each of the six layers we consider (SSU3, SSU2, SSU1, TLS, TTT, and TLT).

Our calculation relies on the vertical profile of atmospheric density from the U.S. standard atmosphere and on the publicly available values of the weighting functions for the three SSU and three MSU layers. The mass weights $\beta(p)$ are defined as follows for each of the N_p layers:

$$\beta(p) = \int_{z_p(\mathrm{BOT})}^{z_p(\mathrm{TOP})} \rho(z)\, V(p,z)\, \Delta(z) \qquad p = 1, \ldots, N_p. \tag{8}$$

where $\rho(z)$ is the density of the standard atmosphere as a function of the height z (in meters), $V(p,z)$ is the SSU or MSU weighting function for the p^{th} atmospheric layer, Δ(z) is the vertical resolution to which $\rho(z)$ and $V(p,z)$ have been interpolated (z =100 meters here), and $N_p = 6$. The vertical integration is from the height of the lowest layer of the p^{th} weighting function, $z_p(\textsc{bot})$, to the height of the top layer of the p^{th} weighting function, $z_p(\textsc{top})$. Realistic land topography is used in the calculation of the density $\rho(z)$.

For each layer, therefore, $\beta(p)$ is the vertical integration of air density weighted by the SSU or MSU weighting function. We normalize each value of $\beta(p)$ by $\beta(\textsc{tot})$, the sum of the six individual $\beta(p)$ values:

$$\beta(p)' = \beta(p) \,/\, \beta(\textsc{tot}) \qquad p = 1, \ldots, N_p. \tag{9}$$

where the $'$ denotes a normalized quantity.

The values of the normalized mass weights (expressed as percentages of the total atmospheric mass sampled by the six sounding channels) are listed below:

1	SSU3	= 0.4%
2	SSU2	= 0.9%
3	SSU1	= 2.1%
4	TLS	= 6.6%
5	TTT	= 39.4%
6	TLT	= 50.6%

In the case of "no mass weighting" shown in Figs. 5 and 6 of the main text and in Figs. S3-S8, all input model and observational latitude-height temperature data sets are multiplied by $\sqrt{W(x)}$, the square root of the area weights for each latitude band. In the "mass weighting" case in Fig. S7, all input temperature data sets are multiplied by $\gamma(x,p)$, the square root of the combined area and mass weights:

$$\gamma(x,p) = \sqrt{W(x)\, \beta(p)'} \qquad x = 1, \ldots, N_x; \quad p = 1, \ldots, N_p \tag{10}$$

The three SSU layers, therefore, sample less than 3.5% of the total mass of the atmosphere. Weighting all input model and observed data sets with the atmospheric mass sampled by individual SSU and MSU layers markedly damps the influence of stratospheric cooling and emphasizes tropospheric warming. In a mass-weighted fingerprint analysis of the SSU+MSU domain, signal strength decreases, noise is amplified, and S/N is reduced by a factor of roughly 4 relative to the case of uniform vertical weights (see Fig. S7).This reduction in S/N is due to multiple factors: the down-weighting of the large global-mean cooling signals in the three SSU channels and TLS, and the reduced impact of the quasi-orthogonality between the signal and noise patterns in the S_{25-50} layer (Fig. S2).

Despite this large reduction in S/N, the mass-weighted fingerprints are still identifiable at the 1% level in each of the 32 individual CMIP extended historical runs and in each of the three observational data sets (Fig. S7).

Weighting function overlap. In all four atmospheric domains considered here (TROP, MSU, SSU, and SSU+MSU; see SI section "Fingerprint analysis") there is overlap between the individual weighting functions used to sample atmospheric temperature changes (22). This overlap can introduce correlation between temperature changes in different atmospheric layers. Of particular concern here is the question of whether S/N results for the six-layer SSU+MSU domain are biased by our use of TTT and TLT (which provide overlapping information about tropospheric temperature change) and by our inclusion of three SSU layers (which provide overlapping information about temperature change in the mid- to upper stratosphere).

We address this question by performing a sensitivity test in which the fingerprint analysis is repeated with three layers only: SSU3, TLS, and TLT. Our choice of these three layers reduces the substantial overlap between weighting functions in the six-layer SSU+MSU case. We refer to the three-layer reduced-space representation of signal, noise, and observations as RED, and we compare fingerprint results in the RED and SSU+MSU cases. This comparison is performed without removal of the global-mean temperature changes in individual atmospheric layers and without any mass weighting of individual layers (see SI sections on "Removal of spatial means" and "Mass and area weighting", respectively).

Results are given in Fig. S8. Relative to the SSU+MSU case, RED systematically reduces signal strength. This reduction occurs because certain signal attributes present in SSU+MSU are absent in RED, such as the amplification of lower tropospheric temperature changes in tropical TTT. Additionally, RED downweights the amplification of cooling in the mid- to upper stratosphere by including results from only one of the three SSU channels used in the six-layer SSU+MSU case.

Figure S8B reveals that the noise amplitude is smaller in RED than in SSU+MSU. This result is partly due to the fact that the noise amplitude is larger in the troposphere than in the stratosphere (see Fig. 5B in the main text). Because RED includes information from only one tropospheric channel (rather than from the two tropospheric channels that are used in SSU+MSU), the noise contribution from the troposphere is smaller in RED than in SSU+MSU.

Additionally, the fingerprint and leading noise modes are spatially more similar in the troposphere than in the mid- to upper stratosphere (compare the TROP and SSU cases in Fig. S2). This pattern similarity contributes to the higher noise in the TROP case in Fig. 5B of the main text – the TROP fingerprint is less successful than the MSU, SSU, and SSU+MSU fingerprints in filtering out internal variability variability. By removing TTT from RED, we are reducing the pattern similarity between tropospheric signal and noise modes, thereby enhancing the effectiveness of noise filtering in RED.

S/N ratios are very similar in the SSU+MSU and RED cases (see Fig. S8C). This similarity occurs because of the compensating effects described above: relative to SSU+MSU, RED has reduced signal strength but also has reduced noise. The RED sensitivity test shows that a simple way of accounting for weighting function overlap – by selectively reducing the number of layers considered in the fingerprint analysis – has a systematic impact on signal and noise, but has relatively little effect on S/N ratios. In both the SSU+MSU and RED cases, S/N ratios by the end of the full 37-year analysis period (1986 to 2022) invariably exceed 35. This holds for fingerprint identification in the three satellite data sets and in all 32 individual CMIP6 HIST$_{ext}$ realizations. We conclude, therefore, that the SSU+MSU fingerprint results presented in the main text are unlikely to be biased by weighting function overlap.

Other statistical analysis details. The sampling distributions of unforced trends in atmospheric temperature shown in Figs. 2 and 3 of the main text were calculated from non-overlapping 37-year and 25-year chunks (respectively) of the same nine CMIP6 pre-industrial control runs used in the fingerprint analysis (see Table S2). While the fingerprint analysis used only 450 years of each control run to ensure that S/N ratios were not biased by models with longer control runs (see SI section "Fingerprint analysis"), the control run trend distributions in Figs. 2 and 3 of the main text were generated using the full length of each control run. The reason for this decision is that unlike in the fingerprint analysis, the "no signal" trend distributions in Figs. 2 and 3 are not being used for statistical significance testing: their primary use is simply to provide visual information regarding differences in the magnitude of forced and unforced trends.

The histograms in Figs. 2 and 3 were plotted with the Matplotlib pyplot.hist function with arrays of weights and with the "density=True" option. This option ensures that "each bin will display the bin's raw count divided by the total number of counts and the bin width... so that the area under the histogram integrates to 1".[¶] The array of weights is defined as:

$$w(j,k) = 1/N_{chunk}(j) \qquad j = 1,\ldots,N_{ctl}; \quad k = 1,\ldots,N_{chunk}(j) \tag{11}$$

where j is an index over the number of pre-industrial control runs, k is an index over the number of non-overlapping 37-year or 25-year least-squares linear trends, and $N_{chunk}(j)$ is the total number of non-overlapping 37-year or 25-year least-squares linear trends in the j^{th} control run.

[¶] https://matplotlib.org/3.3.3/api/_as_gen/matplotlib.pyplot.hist.html

Table S1. Basic information relating to the start dates, end dates, and lengths (N_m, in months) of the 32 CMIP6 historical and SSP5-8.5 simulations used in this study. EM is the "ensemble member" identifier.

	Model	EM	HIST Start	HIST End	HIST N_m	SSP5-8.5 Start	SSP5-8.5 End	SSP5-8.5 N_m
1-2	CESM2	r10i1p1f1, r11i1p1f1	1850-01	2014-12	1980	2015-01	2100-12	1032
3-5	CESM2-WACCM	r1i1p1f1-r3i1p1f1	1850-01	2014-12	1980	2015-01	2100-12	1032
6-8	HadGEM3-GC31-LL	r1i1p1f3–r3i1p1f3	1850-01	2014-12	1980	2015-01	2100-12	1032
9	IPSL-CM6A-LR	r1i1p1f1	1950-01	2014-12	780	2015-01	2300-12	3432
10	IPSL-CM6A-LR	r2i1p1f1	1950-01	2014-12	780	2015-01	2100-12	1032
11-12	IPSL-CM6A-LR	r3i1p1f1, r4i1p1f1	1950-01	2014-12	780	2015-01	2054-12	480
13	IPSL-CM6A-LR	r6i1p1f1	1950-01	2014-12	780	2015-01	2100-12	1032
14	MIROC-ES2L	r1i1p1f2	1850-01	2014-12	1980	2015-01	2100-12	1032
15-16	MPI-ESM-1.2-HR	r1i1p1f1, r2i1p1f1	1850-01	2014-12	1980	2015-01	2100-12	1032
17-26	MPI-ESM-1.2-LR	r1i1p1f1–r10i1p1f1	1850-01	2014-12	1980	2015-01	2100-12	1032
27	MRI-ESM2.0	r1i1p1f1	1850-01	2014-12	1980	2015-01	2300-12	3432
28-31	UKESM1.0-LL	r1i1p1f2–r4i1p1f2	1850-01	2014-12	1980	2015-01	2100-12	1032
32	UKESM1.0-LL	r8i1p1f2	1850-01	2014-12	1980	2015-01	2100-12	1032

Table S2. Start dates, end dates, and lengths (N_m, in months) of the nine CMIP6 pre-industrial control runs used in this study. EM is the "ensemble member" identifier.

	Model	EM	Start	End	N_m
1	CESM2	r1i1p1f1	1-01	1301-12	14400
2	CESM2-WACCM	r1i1p1f1	1-01	499-12	5988
3	HadGEM3-GC31-LL	r1i1p1f1	1850-01	2349-12	6000
4	IPSL-CM6A-LR	r1i1p1f1	1850-01	3049-12	14400
5	MIROC-ES2L	r1i1p1f2	1850-01	2349-12	6000
6	MPI-ESM-1.2-HR	r1i1p1f1	1850-01	2349-12	6000
7	MPI-ESM-1.2-LR	r1i1p1f1	1850-01	2849-12	12000
8	MRI-ESM2.0	r1i1p1f1	1850-01	2550-12	8412
9	UKESM1.0-LL	r1i1p1f2	1960-01	2709-12	9000

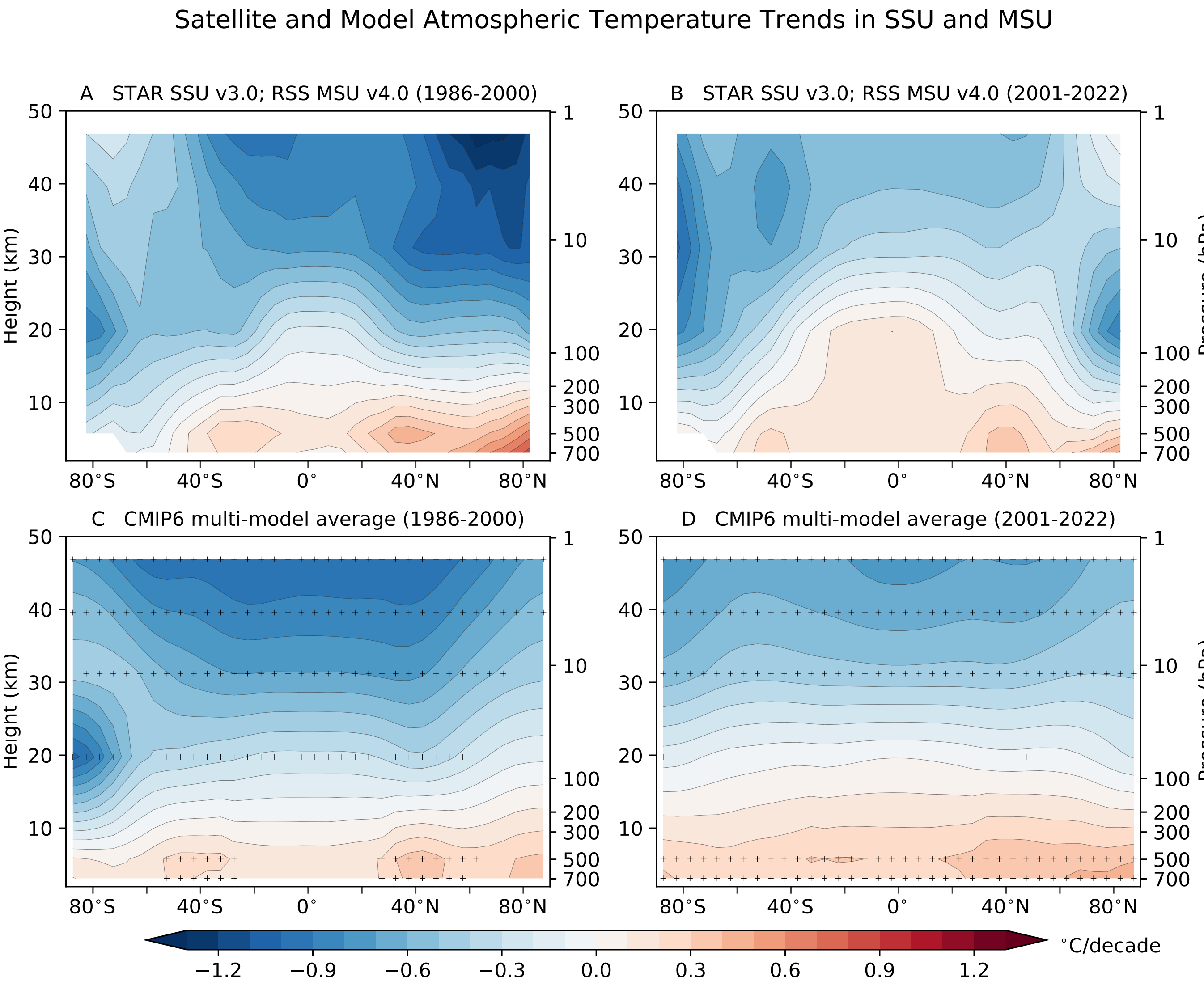

Fig. S1. Trends in zonal-mean annual-mean atmospheric temperature in satellite data and observations. Results are least-squares linear trends over 1986 to 2000 (left column) and over 2001 to 2022 (right column). These two periods are characterized (respectively) by depletion and recovery of observed lower stratospheric ozone concentrations over Antarctica (5, 37, 38). The earlier period is also affected by recovery from the large stratospheric warming signal caused by the 1991 eruption of Pinatubo (see Figs. 1A-D in main text). Observations (panels A, B) are from STAR for the three SSU channels (SSU3, SSU2, and SSU1) (5) and from RSS for MSU TLS, TTT, and TLT (1). Model results (panels C, D) are the multi-model average synthetic SSU and MSU atmospheric temperature trends calculated from 32 realizations of HIST$_{ext}$ runs performed with nine different CMIP6 models. In all panels, global-mean temperature changes are retained for each of the six atmospheric layers considered. The black dots in panels C and D denote latitude bands and layers with local S/N ratios ≥ 2: i.e., locations where the multi-model average trend over the analysis period is at least a factor of two larger than the standard deviation of individual model trends. Black dots are plotted at the approximate peaks of the three SSU and three MSU weighting functions.

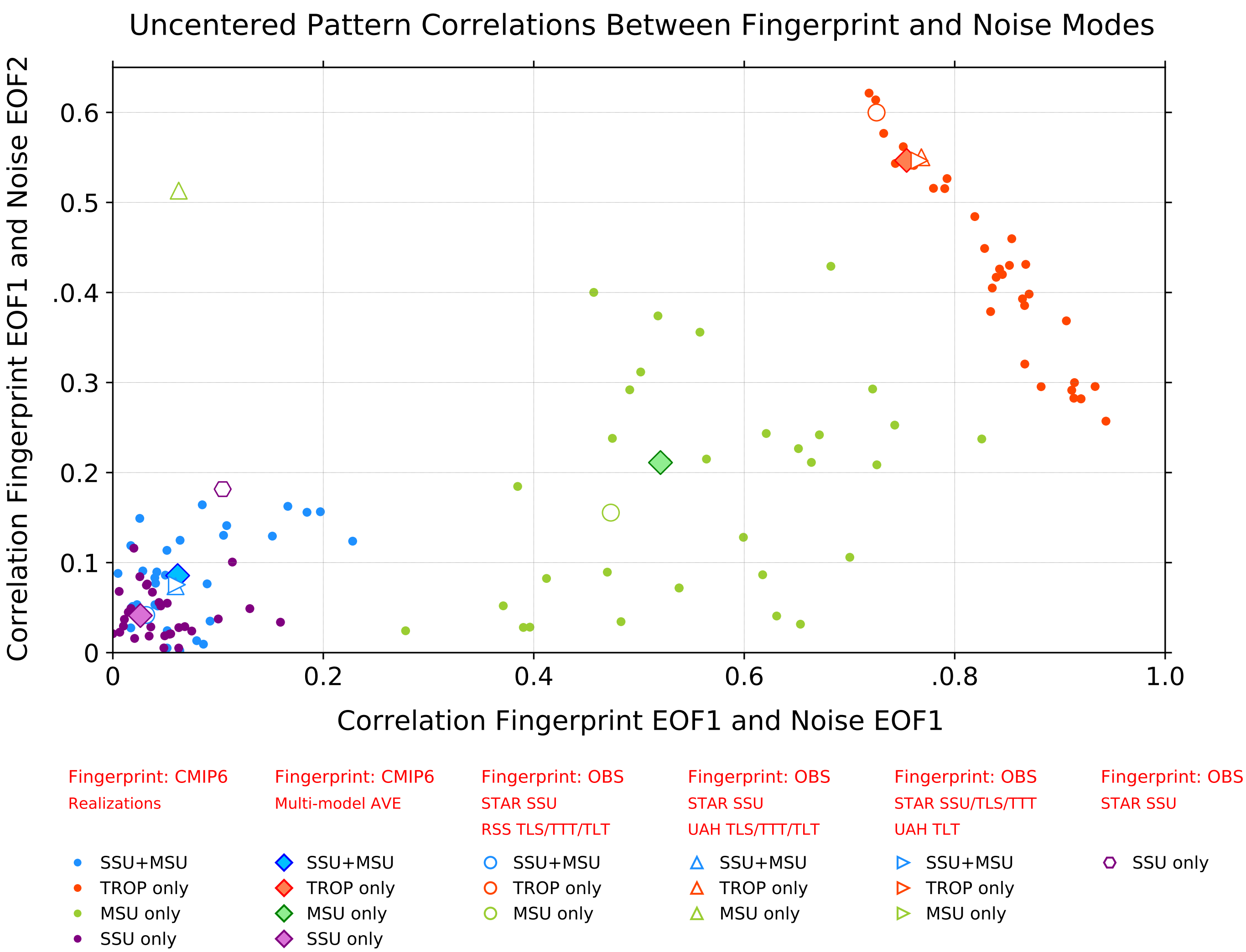

Fig. S2. Values of the uncentered pattern correlations between the fingerprint F and the first two noise modes in CMIP6 simulations. Results are for four spatial domains: SSU+MSU, TROP, MSU, and SSU. For each domain, F was estimated from three sources: the 32 individual model HIST$_{ext}$ realizations performed with 9 different CMIP6 models (filled circles), the multi-model average HIST$_{ext}$ atmospheric temperature changes (filled diamonds), and the satellite data (unfilled symbols). The first two noise Empirical Orthogonal Functions (EOFs) were calculated using 4,050 years of concatenated pre-industrial control run data. Pattern correlations between F and noise EOFs 1 and 2 are plotted on the x-axis and y-axis (respectively). Noise EOFs 1 and 2 are shown in the middle and right columns of Fig. 6 of the main text; the fingerprints estimated from the CMIP6 multi-model average HIST$_{ext}$ data are in the left column of Fig. 6. For the SSU+MSU domain, the F patterns for selected individual HIST$_{ext}$ realizations are displayed in Figs. S3A-I and the F patterns for the two satellite data sets are given in Figs. S3K and L. In calculating fingerprints and noise EOFs, global-mean temperature changes were retained for each of the six atmospheric layers considered. The data used for computing EOFs were area-weighted but not mass-weighted. Since the signs of the fingerprints and noise EOFs are arbitrary, we show the absolute value of the pattern correlation.

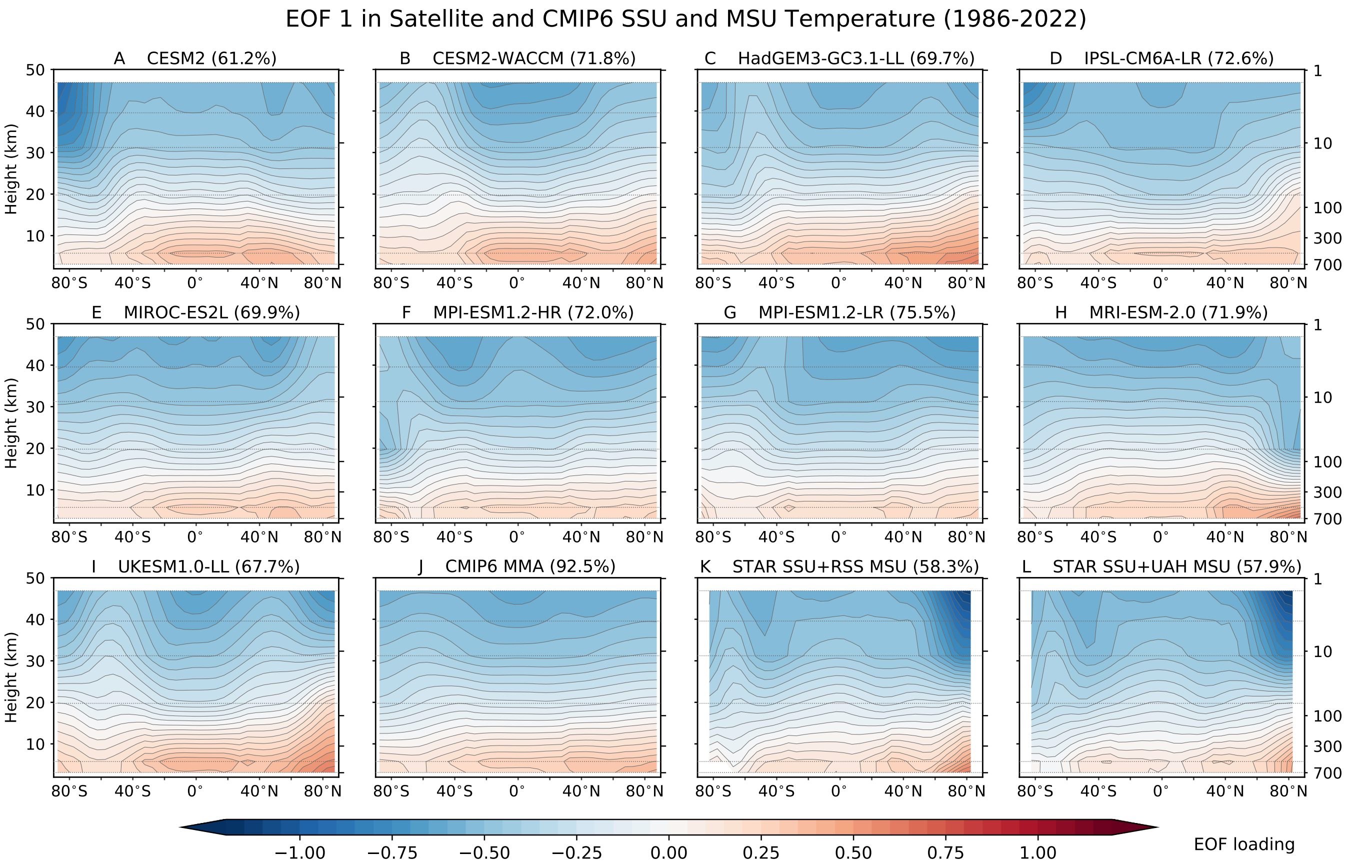

Fig. S3. Fingerprint pattern of zonal-mean annual-mean atmospheric temperature change in simulations and observations for the SSU+MSU domain. Results are the first Empirical Orthogonal Function (EOF) of HIST$_{ext}$ simulations in individual CMIP6 models (panels A-I) and in the CMIP6 multi-model average (panel J). The leading EOF for two satellite data sets is also shown (panels K, L). EOFs are calculated over 1986 to 2022 using temperature changes for six atmospheric layers (SSU3, SSU2, SSU1, TLS, TTT, and TLT). For models with multiple HIST$_{ext}$ realizations in panels A-I, results are for the first realization only. In all EOF calculations, global-mean temperature changes are retained for each of the six atmospheric layers considered. The dotted horizontal grey lines are plotted at the approximate peaks of the three SSU and three MSU weighting functions. The explained variance of each EOF is indicated in the panel title (in parentheses). The data used for computing EOFs were area-weighted but not mass-weighted.

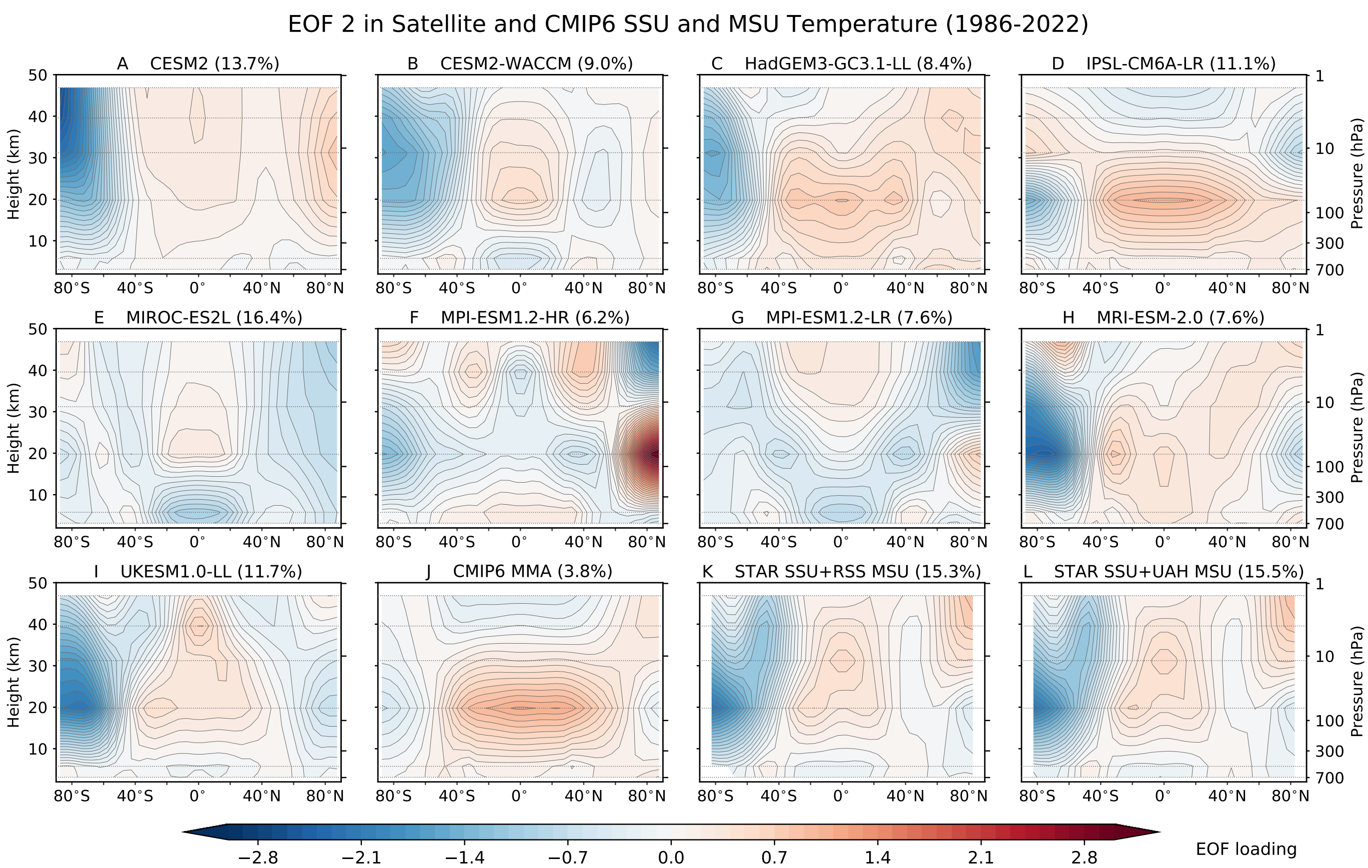

Fig. S4. As for Fig. S3 but for EOF 2.

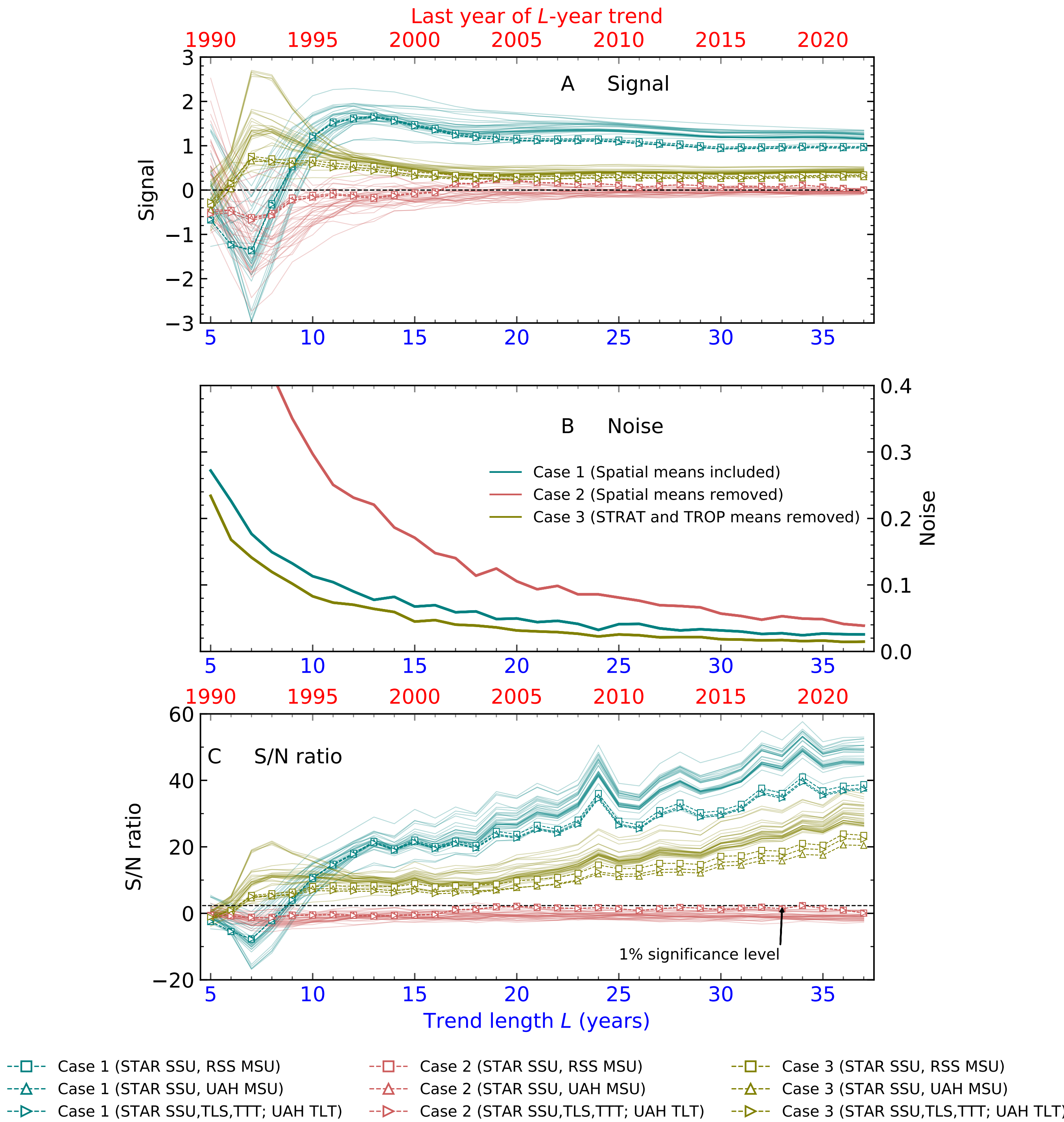

Fig. S5. Signal, noise, and S/N ratios (panels A-C, respectively) in model and observational SSU and MSU data. Results are for the six-layer SSU+MSU case (SI section "Fingerprint analysis"). The latitude-height temperature changes for these six layers are used in three sets of calculations. In Case 1, the global-mean temperature change over time is retained in each layer. In Case 2, each layer's global-mean temperature-change is removed. Case 3 is similar to Case 2, but involves subtraction of the stratospheric-average global-mean change from each individual stratospheric layer and the tropospheric-average global-mean change from each individual tropospheric layer (see SI section "Removal of spatial means"). As in Figs. 5A and C of the main text, all signals and S/N ratios in which observed temperature data are used for signal calculation are plotted with symbols and dashed lines. "Model only" results are plotted with solid lines. The dashed horizontal line in panel C is the 1% significance level. The data used for computing EOFs were area-weighted but not mass-weighted.

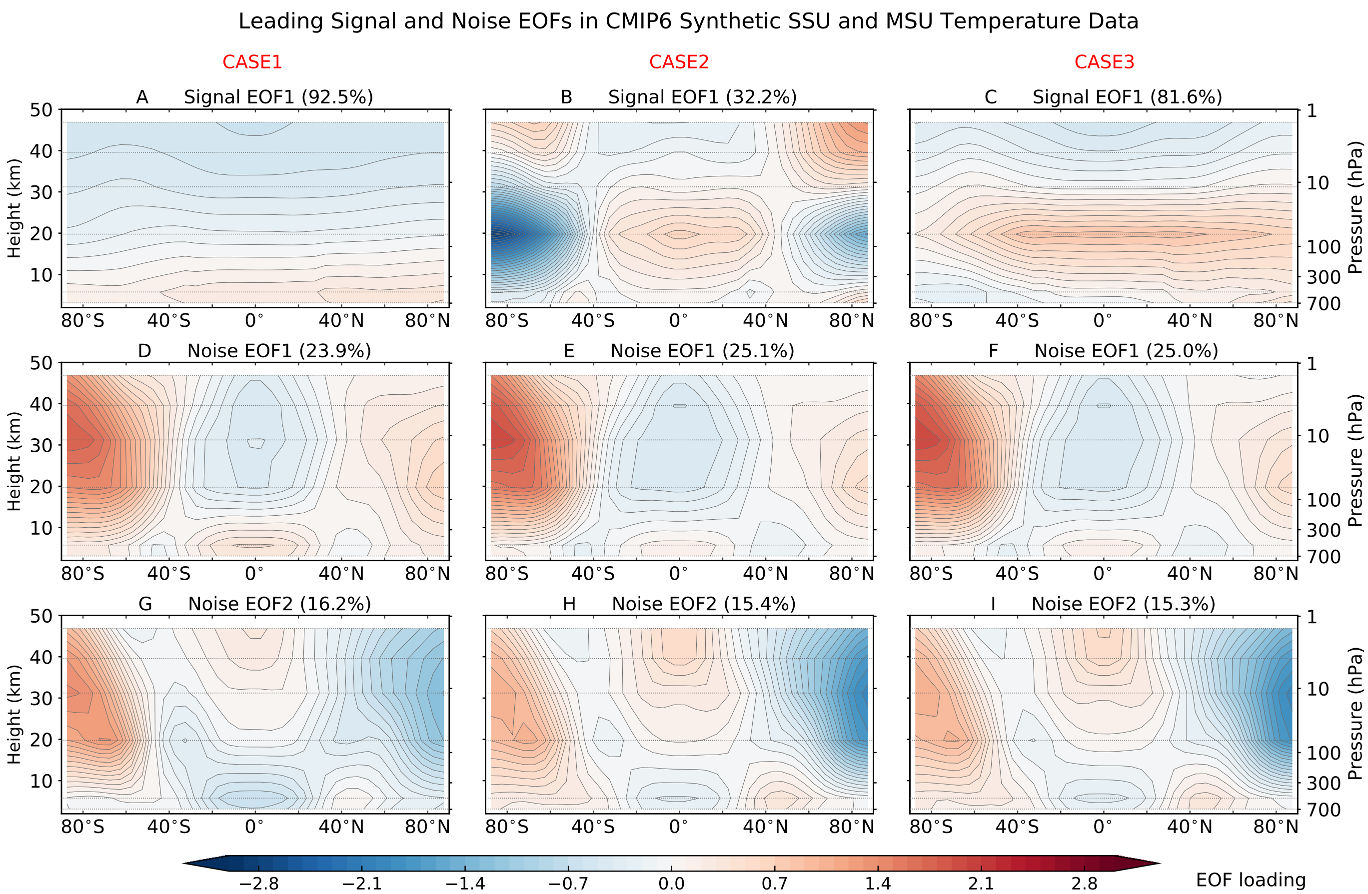

Fig. S6. Fingerprints and leading noise modes in CMIP6 simulations. Results are for the SSU+MSU domain. The fingerprint (row 1) is EOF 1 of the multi-model average atmospheric temperature changes computed from 32 realizations of HIST$_{ext}$ runs performed with nine CMIP6 models. The first two noise EOFs (rows 2 and 3) were calculated from concatenated pre-industrial control runs with the same nine models. Fingerprints and noise EOFs are for Cases 1, 2, and 3 (columns 1-3). These three cases consider the impact of different decisions regarding removing or retaining global-mean temperature changes (see SI section "Removal of spatial means"). The data used for computing EOFs were area-weighted but not mass-weighted. The dotted horizontal gray lines are plotted at the approximate peaks of the SSU and MSU weighting functions. The noise modes in Cases 1, 2, and 3 are highly similar because their patterns are dominated by variability at smaller spatial scales, and are therefore relatively unaffected by removal or inclusion of the global-mean temperature changes in Cases 2 and 3. The prominent latitudinally coherent maximum at TLS level in panel C is due to the fact that the global-mean cooling of TLS over 1986 to 2022 is at least a factor of three smaller than the global-mean cooling in the three SSU channels (see Fig. 2 in the main text).

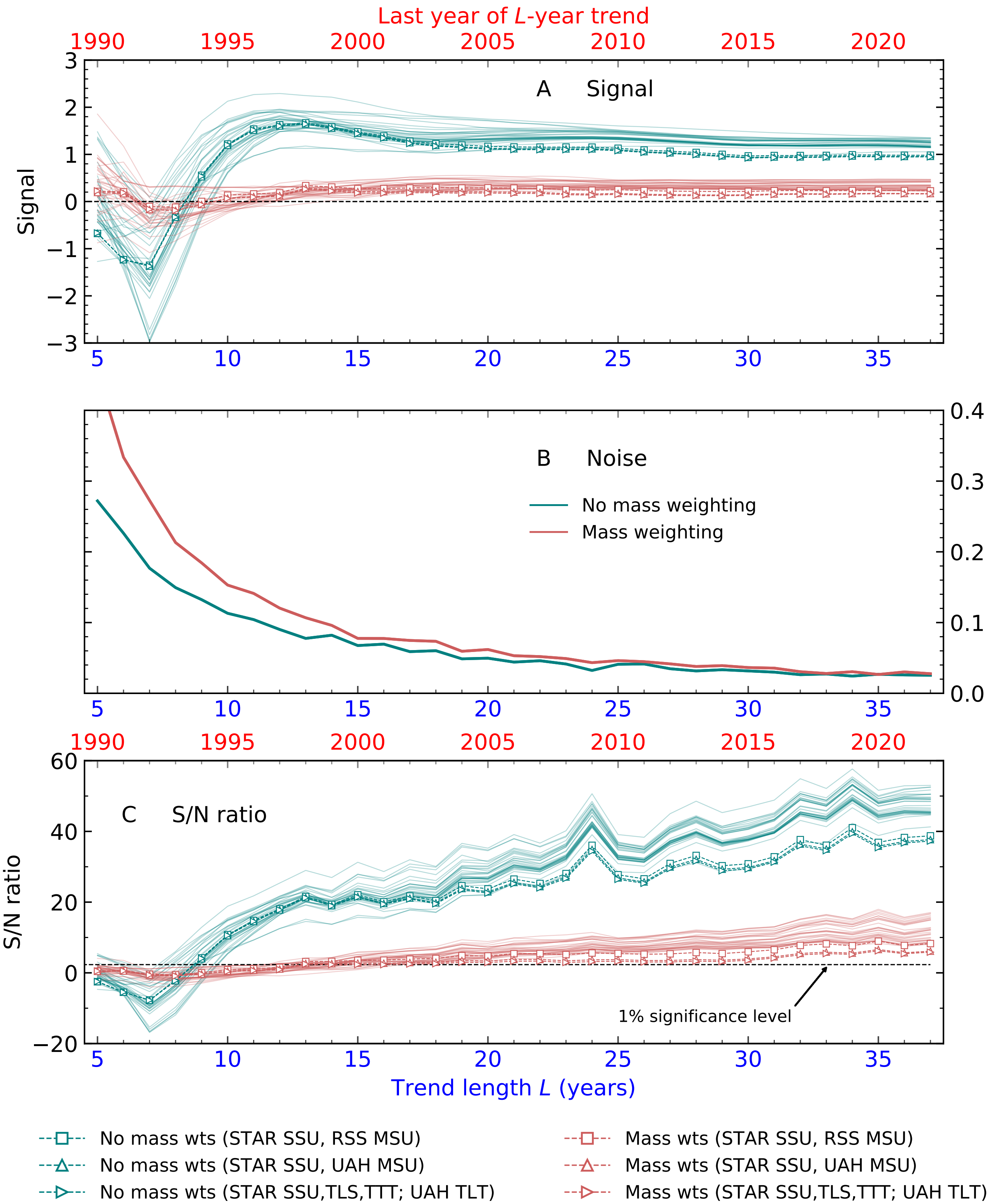

Fig. S7. Sensitivity of signal, noise, and S/N ratios to vertical weighting (panels A-C, respectively). Results are for the six-layer SSU+MSU domain; the global-mean temperature changes are included for each layer. The annual-mean latitude-height temperature changes for these six layers are used in two different sets of calculations. In the "no mass weighting" case, each of the six individual layers is given equal weight in the fingerprint analysis. Results for this case are identical to the results shown for the SSU+MSU case in Fig. 5 of the main text. In the "mass weighting" case, weights representative of the atmospheric mass sampled by each of the SSU and MSU weighting functions are applied to the temperature changes in each layer (see SI section "Mass and area weighting"). Mass weighting is performed for each model and observational data set. As in Figs. 5A and C of the main text, all signals and S/N ratios in which observed temperature data are used for signal calculation are plotted with symbols and dashed lines. "Model only" results are plotted with solid lines. The dashed horizontal line in panel C is the 1% significance level.

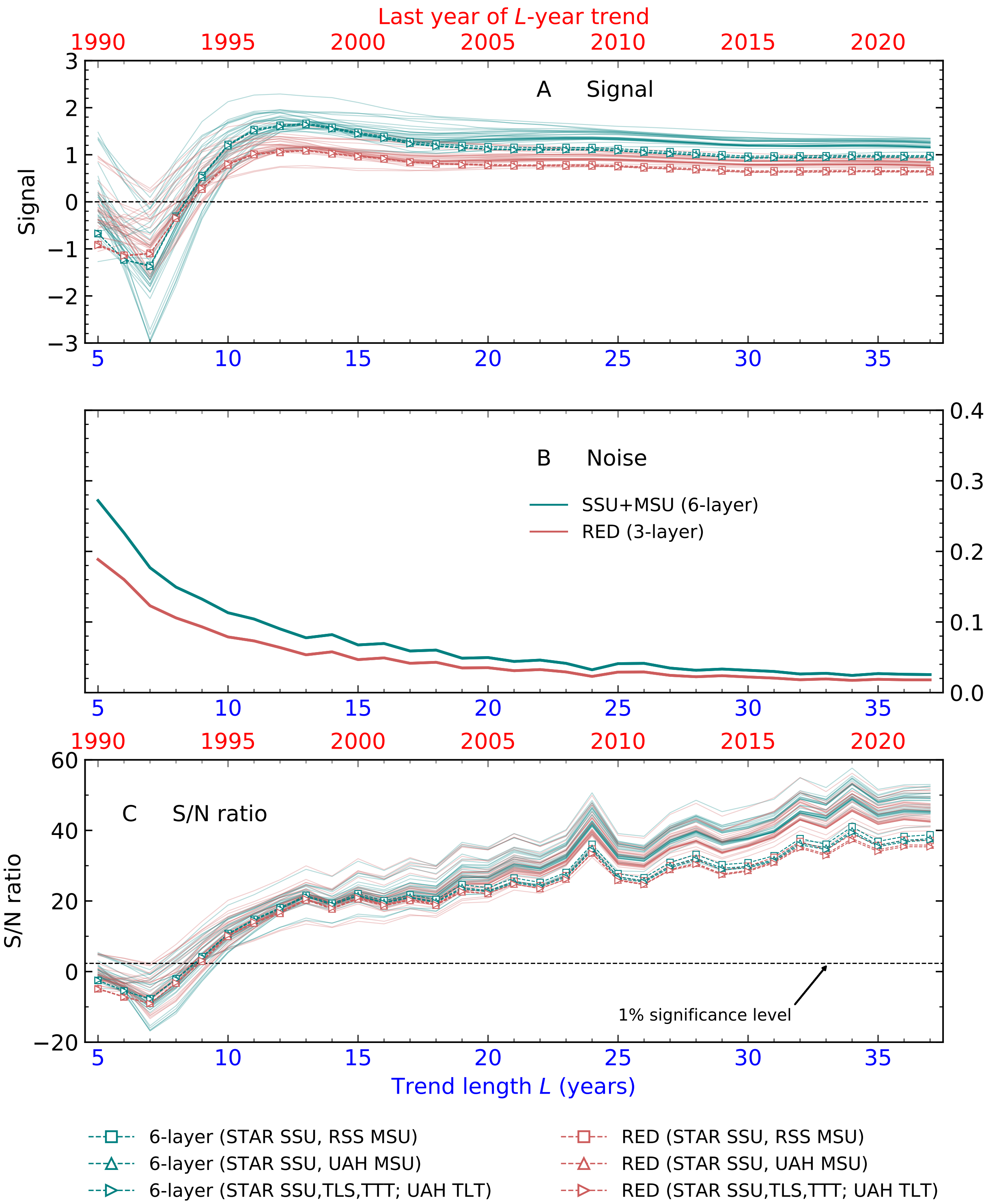

Fig. S8. Sensitivity of signal, noise, and S/N ratios to the degree of overlap between weighting functions (panels A-C, respectively). Results are for two different cases: SSU+MSU and RED. SSU+MSU comprises annual-mean latitude-height temperature-change information from six atmospheric layers (the three SSU channels and MSU TLS, TTT, and TLT). There is substantial overlap between the weighting functions for these six layers (22), leading to overlap in the portions of the atmosphere that the weighting functions sample. RED reduces this overlap by using information from three selected layers only: SSU3, TLS, and TLT (see SI section "Weighting function overlap"). Both SSU+MSU and RED include global-mean temperature changes for each layer considered. As in Figs. 5A and C of the main text, all signals and S/N ratios in which observed temperature data are used for signal calculation are plotted with symbols and dashed lines. "Model only" results are plotted with solid lines. The dashed horizontal line in panel C is the 1% significance level. The data used for computing EOFs were area-weighted but not mass-weighted.

References

1. C Mears, FJ Wentz, Sensitivity of satellite-derived tropospheric temperature trends to the diurnal cycle adjustment. *J. Clim.* **29**, 3629–3646 (2016).
2. RW Spencer, JR Christy, WD Braswell, UAH version 6 global satellite temperature products: Methodology and results. *Asia-Pac. J. Atmos. Sci.* **53**, 121–130 (2017).
3. CZ Zou, MD Goldberg, X Hao, New generation of U.S. satellite microwave sounder achieves high radiometric stability performance for reliable climate change detection. *Sci. Adv.* **4**, eaau0049 (2018).
4. CZ Zou, H Xu, X Hao, Q Liu, Mid-tropospheric layer temperature record derived from satellite microwave sounder observations with backward merging approach. *J. Geophys. Res.* **128**, e2022JD037472 (2023).
5. CZ Zou, H Qian, Stratospheric temperature climate record from merged SSU and AMSU-A observations. *J. Atmos. Ocean. Tech.* **33**, 1967–1984 (2016).
6. CZ Zou, H Qian, W Wang, L Wang, C Long, Recalibration and merging of SSU observations for stratospheric temperature trend studies. *J. Geophys. Res.* **119**, 13180–13205 (2014).
7. Q Fu, CM Johanson, SG Warren, DJ Seidel, Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature* **429**, 55–58 (2004).
8. V Eyring, et al., Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Mod. Dev.* **9(5)**, 1937–1958 (2016).
9. K Riahi, et al., The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Glob. Env. Chang.* **42**, 153–168 (2017).
10. DWJ Thompson, et al., The mystery of recent stratospheric temperature trends. *Nature* **491**, 692–697 (2012).
11. WJ Collins, et al., AerChemMIP: Quantifying the effects of chemistry and aerosols in CMIP6. *Geosci. Model. Dev. Discuss.* **10**, 585–607 (2017).
12. S Solomon, et al., The persistently variable "background" stratospheric aerosol layer and global climate change. *Science* **333**, 866–870 (2011).
13. BD Santer, et al., Human and natural influences on the changing thermal structure of the atmosphere. *Proc. Nat. Acad. Sci.* **110**, 17235–17240 (2013).
14. BD Santer, et al., Using climate model simulations to constrain observations. *J. Clim.* **34**, 6281–6301 (2021).
15. BD Santer, et al., Comparing tropospheric warming in climate models and satellite data. *J. Clim.* **30**, 3–4 (2017).
16. Q Fu, CM Johanson, Stratospheric influences on MSU-derived tropospheric temperature trends: A direct error analysis. *J. Clim.* **17**, 4636–4640 (2004).
17. Q Fu, CM Johanson, Satellite-derived vertical dependence of tropical tropospheric temperature trends. *Geophys. Res. Lett.* **32**, L10703 (2005).
18. CM Johanson, Q Fu, Robustness of tropospheric temperature trends from MSU Channels 2 and 4. *J. Clim.* **19**, 4234–4242 (2006).
19. Q Fu, S Manabe, CM Johanson, On the warming in the tropical upper troposphere: Models versus observations. *Geophys. Res. Lett.* **38**, L15704 (2011).
20. S Po-Chedley, TJ Thorsen, Q Fu, Removing diurnal cycle contamination in satellite-derived tropospheric temperatures: Understanding tropical tropospheric trend discrepancies. *J. Clim.* **28**, 2274–2290 (2015).
21. S Po-Chedley, et al., Natural variability drives model-observational differences in tropical tropospheric warming. *Proc. Nat. Acad. Sci.* **118**, e2020962118 (2021).
22. A Steiner, et al., Observed temperature changes in the troposphere and stratosphere from 1979 to 2018. *J. Clim.* **33**, 8165–8194 (2020).
23. B Hassler, et al., Comparison of three vertically resolved ozone data sets: climatology, trends and radiative forcings. *Atmos. Chem. Phys.* **13**, 5533–5550 (2013).
24. NP Gillett, BD Santer, AJ Weaver, Quantifying the influence of stratospheric cooling on satellite-derived tropospheric temperature trends. *Nature* **432** (2004).
25. JT Kiehl, J Caron, JJ Hack, On using global climate model simulations to assess the accuracy of MSU retrieval methods for tropospheric warming trends. *J. Clim.* **18**, 2533–2539 (2005).
26. K Hasselmann, *On the signal-to-noise problem in atmospheric response studies.* (Roy. Met. Soc., London), pp. 251–259 (1979).
27. BD Santer, et al., Identifying human influences on atmospheric temperature. *Proc. Nat. Acad. Sci.* **110**, 26–33 (2013).
28. BD Santer, et al., Signal-to-noise analysis of time-dependent greenhouse warming experiments. *Cli. Dyn.* **9**, 267–285 (1994).
29. BD Santer, et al., Human influence on the seasonal cycle of tropospheric temperature. *Science* **361**, eaas8806 (2018).
30. BD Santer, et al., Celebrating the anniversary of three key events in climate change science. *Nat. Clim. Chang.* **9**, 180–182 (2019).
31. S Sippel, N Meinshausen, EM Fischer, E Székely, R Knutti, Climate change now detectable from any single day of weather at global scale. *Nat. Clim. Chang.* **10**, 35–41 (2020).
32. HM van den Dool, S Saha, Å Johansson, Empirical orthogonal teleconnections. *J. Clim.* **13**, 1421–1435 (2000).
33. BD Santer, et al., Quantifying stochastic uncertainty in detection time of human-caused climate signals. *Proc. Nat. Acad. Sci.* **116**, 19821–19827 (2019).

34. BD Santer, et al., Influence of satellite data uncertainties on the detection of externally forced climate change. *Science* **300**, 1280–1284 (2003).
35. S Manabe, RT Wetherald, Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci.* **24**, 241–259 (1967).
36. BD Santer, et al., Amplification of surface temperature trends and variability in the tropical atmosphere. *Science* **309**, 1551–1556 (2005).
37. WJ Randel, AK Smith, F Wu, CZ Zou, H Qian, Stratospheric temperature trends over 1979–2015 derived from combined SSU, MLS, and SABER satellite observations. *J. Clim.* **29**, 4843–4859 (2016).
38. S Solomon, et al., Mirrored changes in Antarctic ozone and stratospheric temperature in the late 20th versus early 21st centuries. *J. Geophys. Res.* **122**, 8940–8950 (2017).