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 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:

STAR1 = STAR SSU3/2/1 + STAR TLS/TTT + RSS TLT
STAR2 = STAR SSU3/2/1 + STAR TLS/TTT + UAH TLT

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:

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

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:

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

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 $5^\circ \times 5^\circ$ 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 $5^\circ \times 5^\circ$ 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 $9 \times 450 = 4{,}050$ 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 $5^\circ \times 5^\circ$ 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:

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

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:

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

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:

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

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:

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

where * denotes departures from the global-mean.

In Case 3, $< T_{obs}\{\text{STRAT}\}(t) >$ and $< T_{obs}\{\text{TROP}\}(t) >$ 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:

$$T_{obs}(x,p,t)^{**} = T_{obs}(x,p,t) - < T_{obs}\{\text{STRAT}\}(t) >$$
$$x = 1, \ldots, N_x; \quad p = 1, \ldots, 4; \quad t = 1, \ldots, 37. \qquad [7]$$
$$T_{obs}(x,p,t)^{**} = T_{obs}(x,p,t) - < T_{obs}\{\text{TROP}\}(t) >$$
$$x = 1, \ldots, N_x; \quad p = 5, 6; \quad t = 1, \ldots, 37.$$

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, $z_p(\mathrm{BOT})$, to the height of the top layer of the p^{th} weighting function, $z_p(\mathrm{TOP})$. 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 $\beta(\mathrm{TOT})$, the sum of the six individual $\beta(p)$ values:

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

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:

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

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

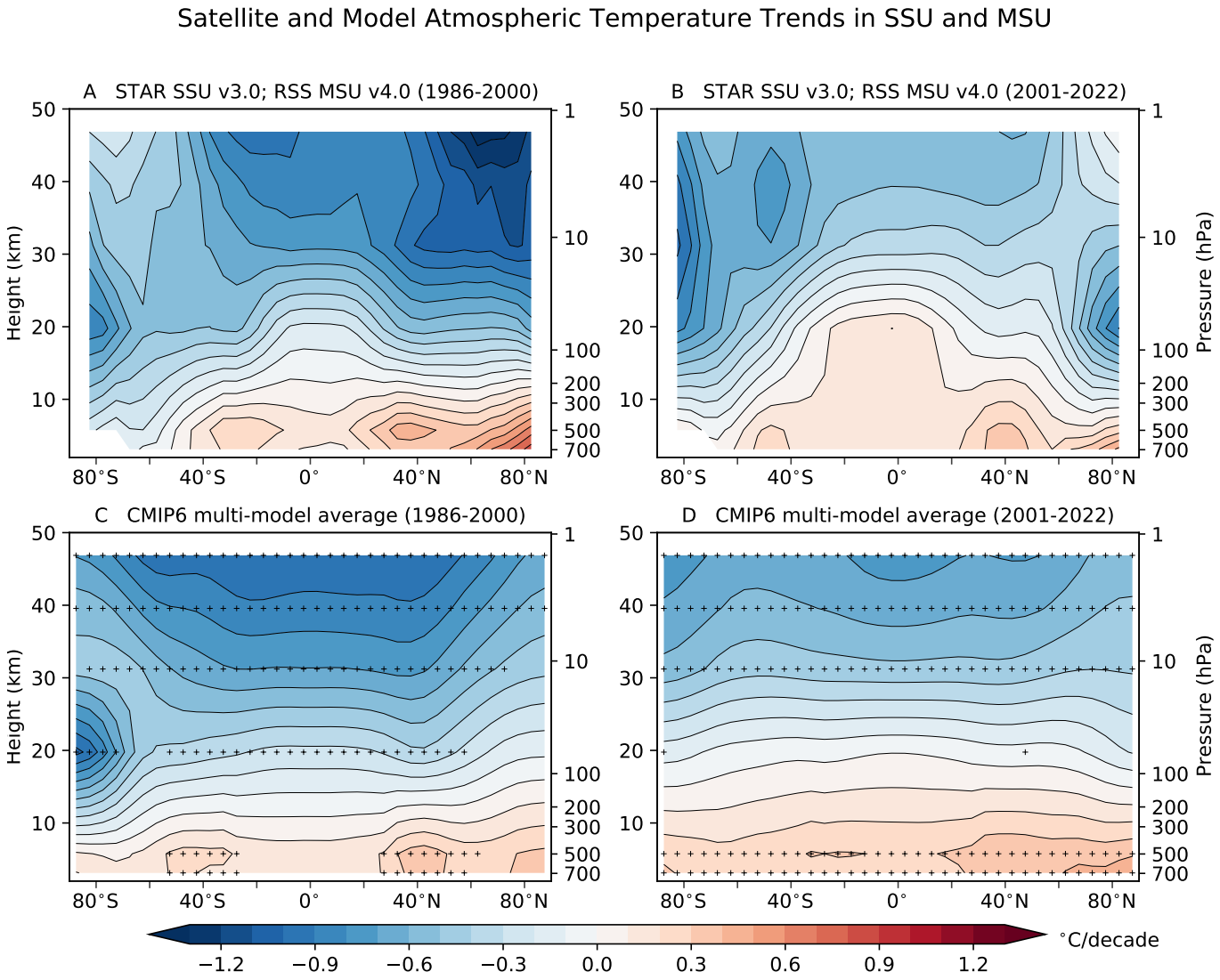

Fig. S1. Trends in zonal-mean annual-mean atmospheric temperature in satellite data and observations. Results are least-squares linear trends over 1986 to 2000 (left column) and over 2001 to 2022 (right column). These two periods are characterized (respectively) by depletion and recovery of observed 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.

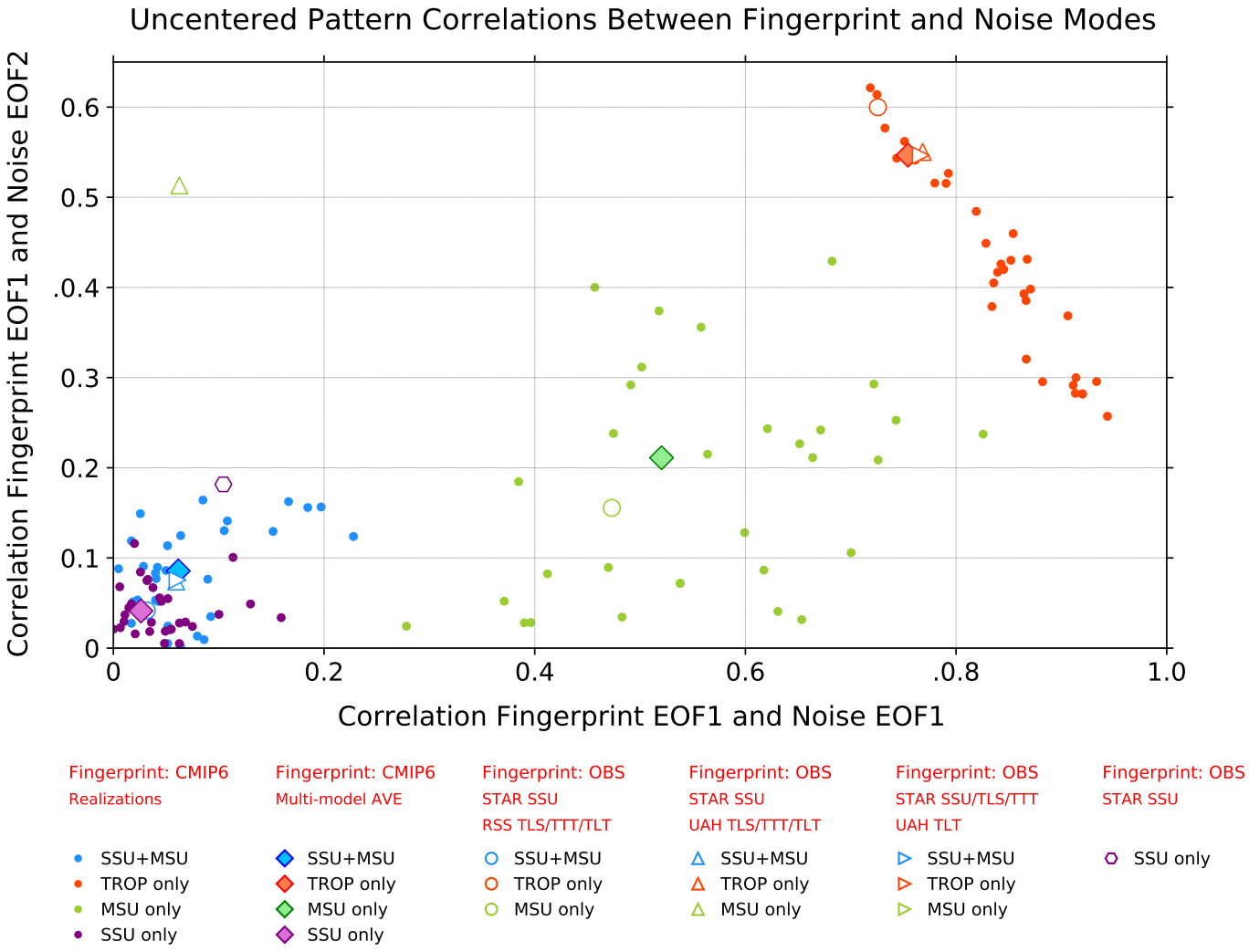

Fig. S2. Values of the uncentered pattern correlations between the fingerprint F and the first two noise modes in CMIP6 simulations. Results are for four spatial domains: SSU+MSU, TROP, MSU, and SSU. For each domain, F was estimated from three sources: the 32 individual model HIST$_{ext}$ realizations 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.

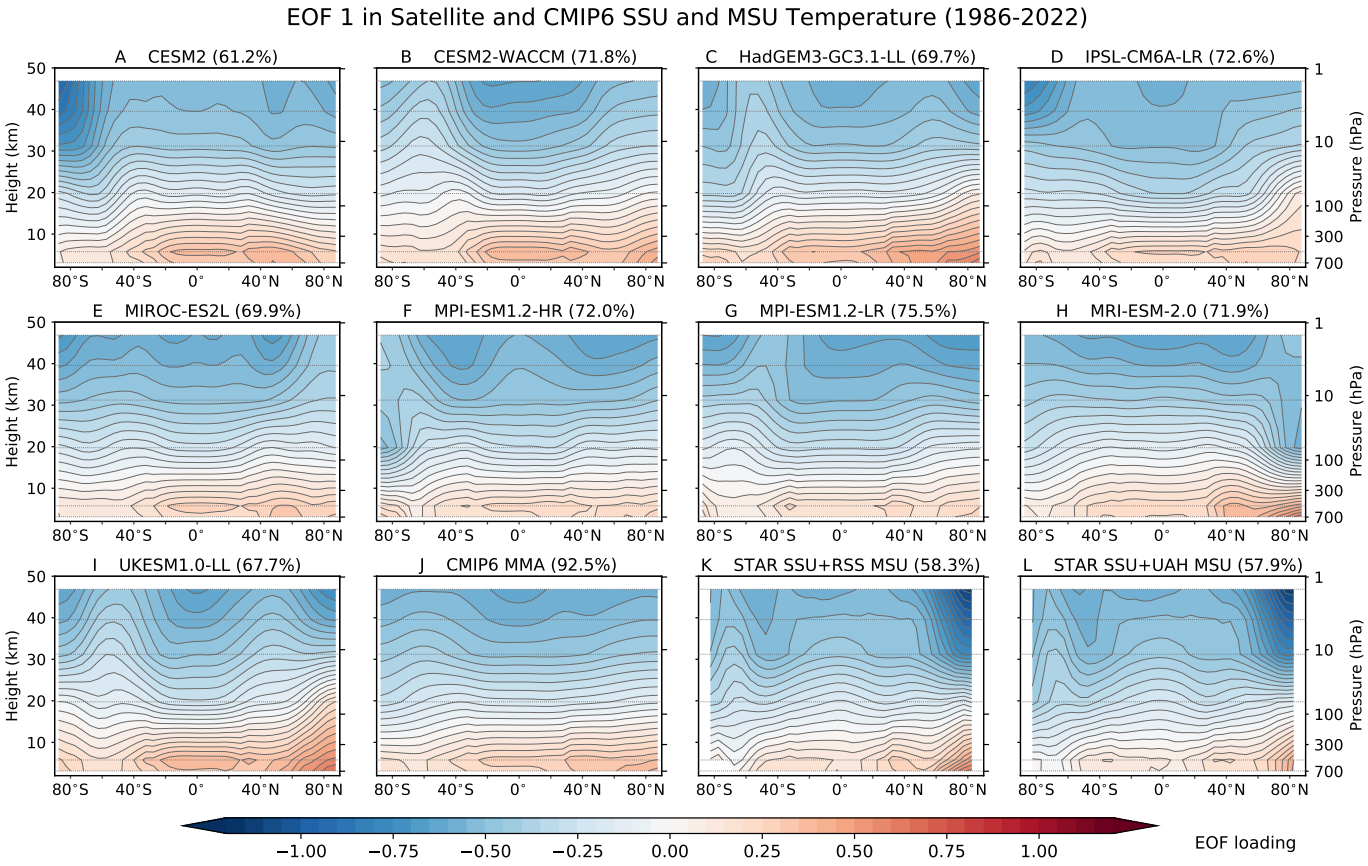

Fig. S3. Fingerprint pattern of zonal-mean annual-mean atmospheric temperature change in simulations and observations for the SSU+MSU domain. Results are the first Empirical Orthogonal Function (EOF) of HIST$_{ext}$ simulations in individual CMIP6 models (panels A-I) and in the CMIP6 multi-model average (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.

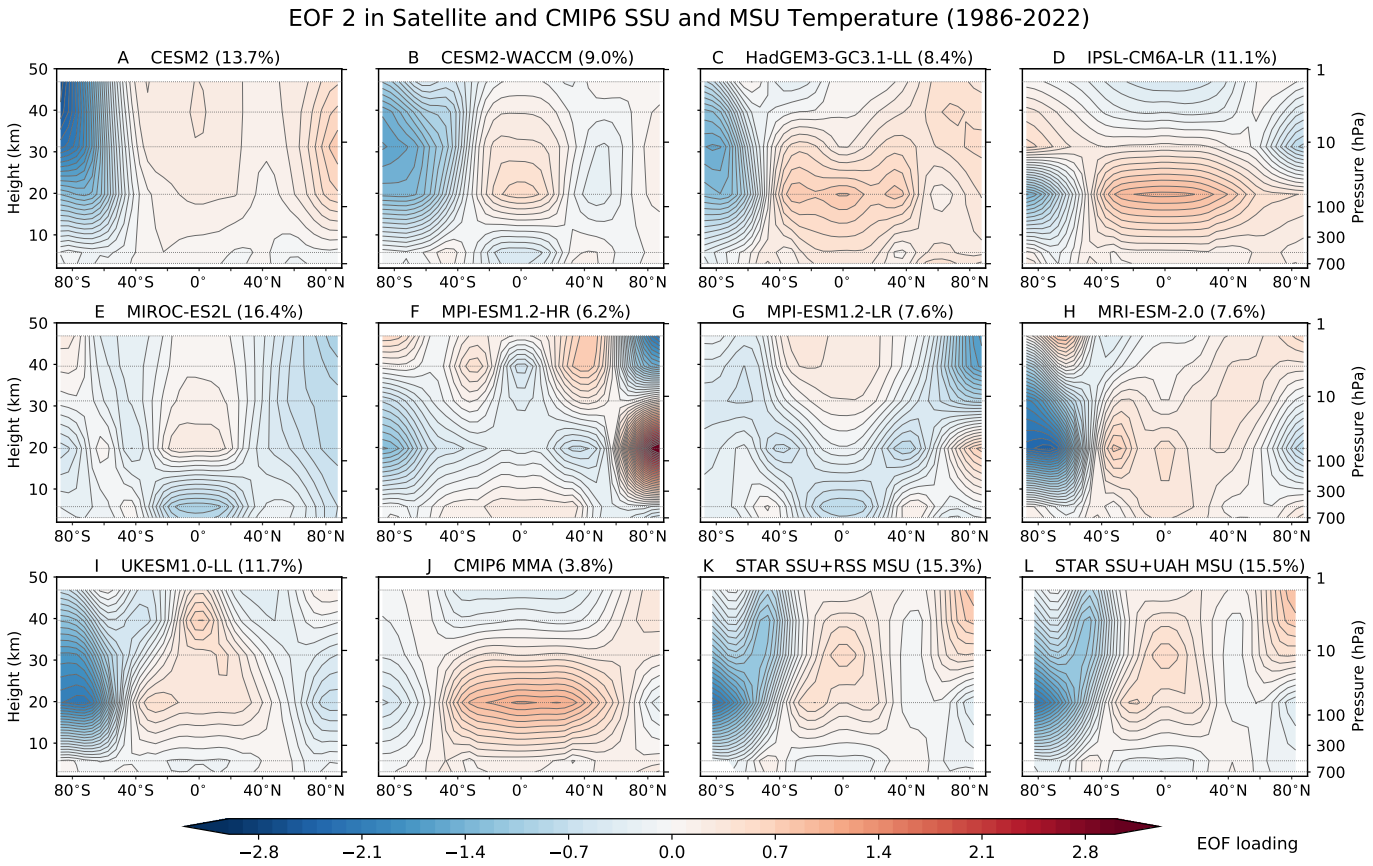

Fig. S4. As for Fig. S3 but for EOF 2.

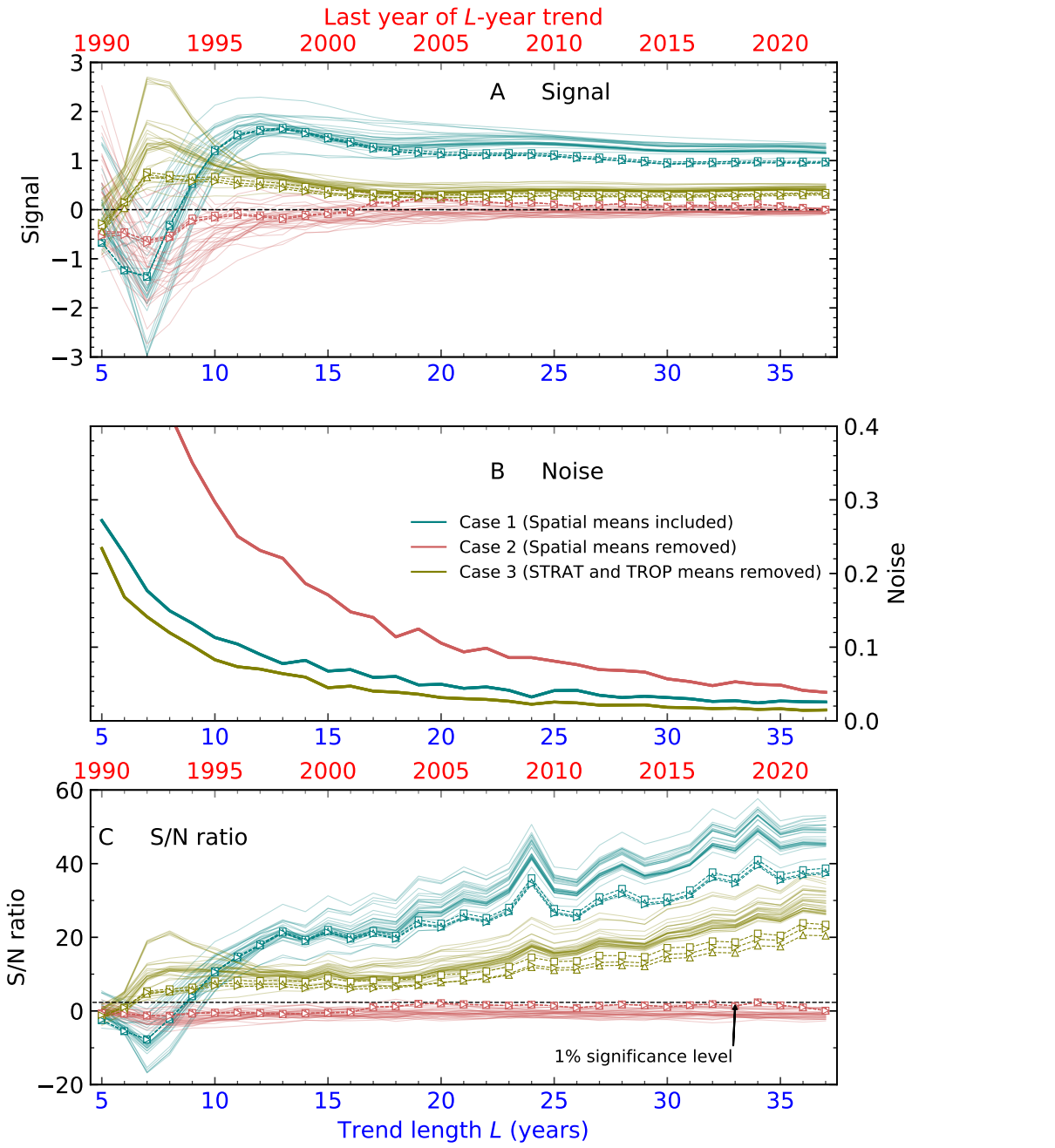

Fig. S5. Signal, noise, and S/N ratios (panels A-C, respectively) in model and observational SSU and MSU data. Results are for the six-layer SSU+MSU case (SI section "Fingerprint analysis"). The latitude-height temperature changes for these six layers are used in three sets of calculations. In Case 1, 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.

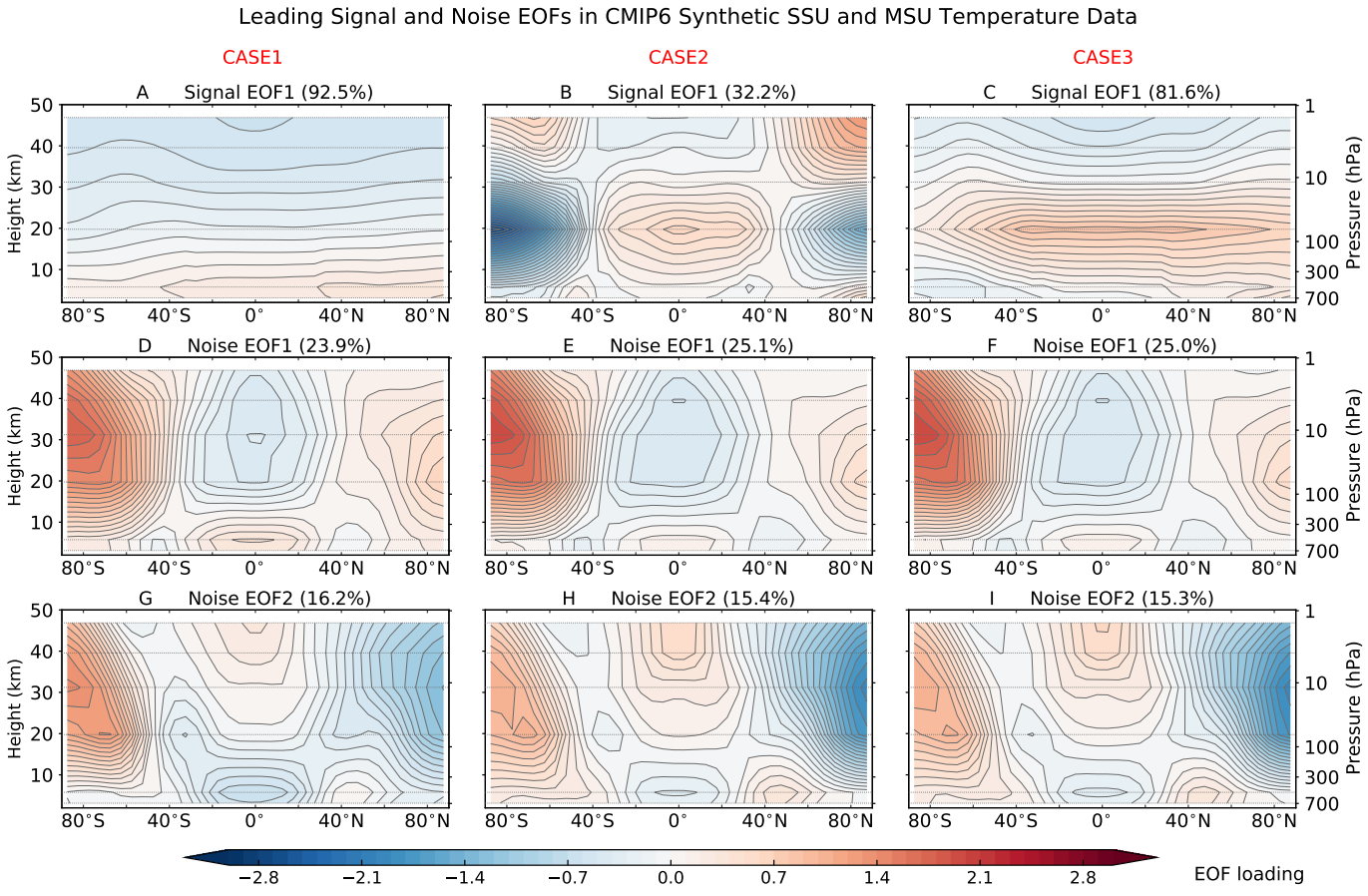

Fig. S6. Fingerprints and leading noise modes in CMIP6 simulations. Results are for the SSU+MSU domain. The fingerprint (row 1) is EOF 1 of the multi-model average atmospheric temperature changes computed from 32 realizations of HIST_{ext} runs performed with nine CMIP6 models. The first two noise EOFs (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).

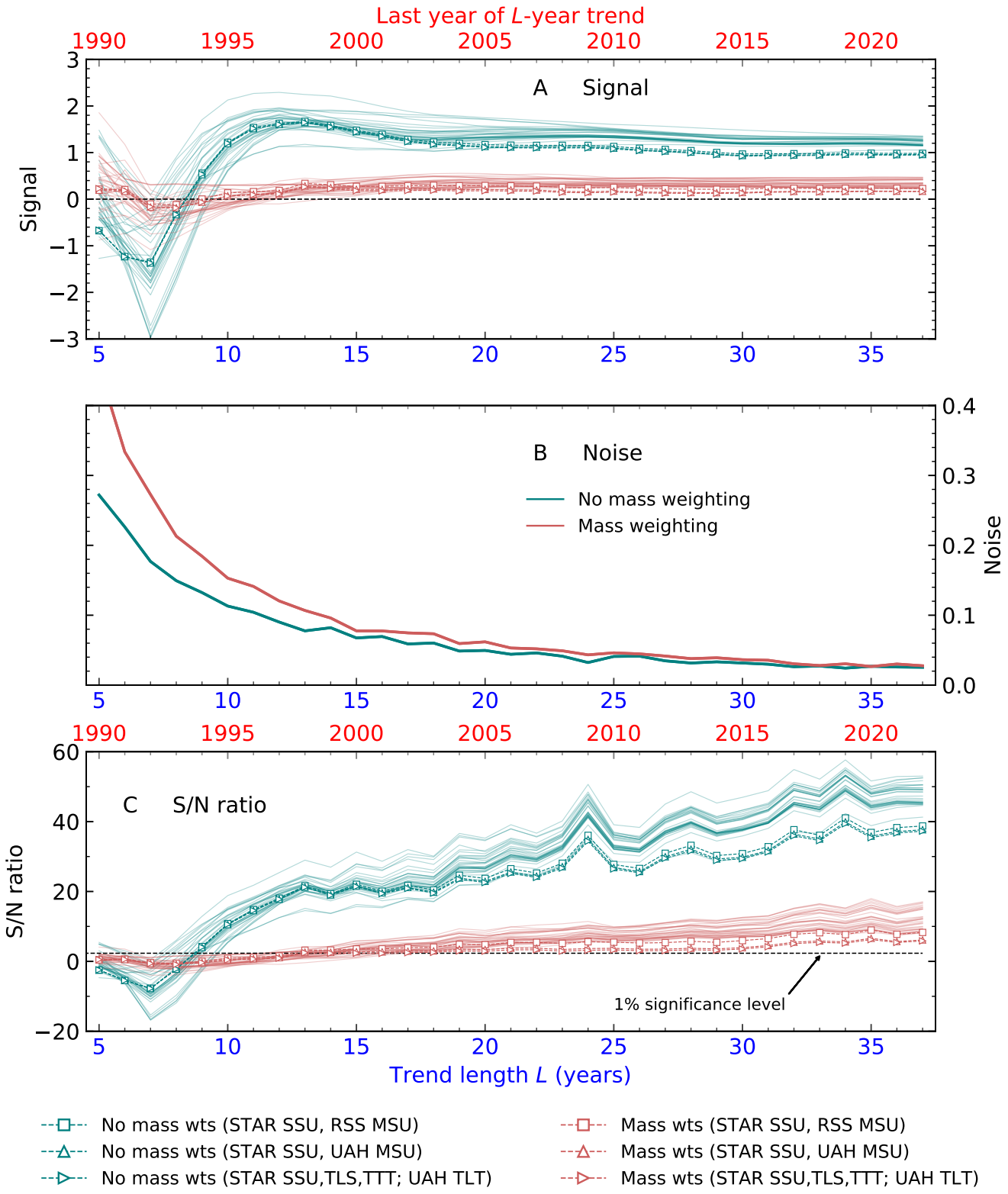

Fig. S7. Sensitivity of signal, noise, and S/N ratios to vertical weighting (panels A-C, respectively). Results are for the six-layer SSU+MSU domain; the global-mean temperature changes are included for each layer. The annual-mean latitude-height temperature changes for these six layers are used in two 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.

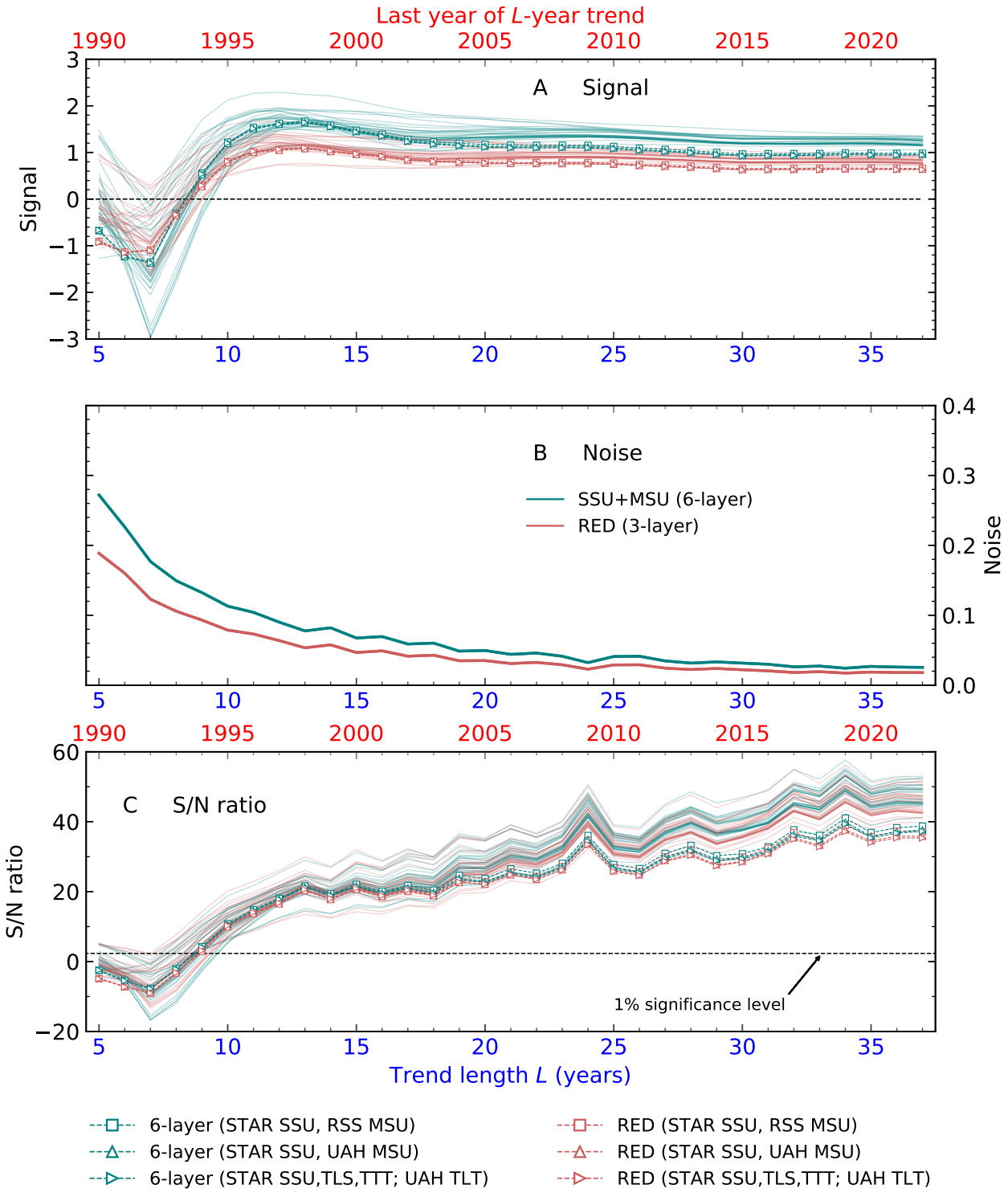

Fig. S8. Sensitivity of signal, noise, and S/N ratios to the degree of overlap between weighting functions (panels A-C, respectively). Results are for two different cases: SSU+MSU and RED. SSU+MSU comprises annual-mean latitude-height temperature-change information from six atmospheric layers (the three 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).