10.1175/JCLI-D-16-0333.s1

Supplemental Material

1 Satellite atmospheric temperature data

Since late 1978, microwave sounders on NOAA polar-orbiting satellites have measured the microwave emissions of oxygen molecules. Because oxygen molecules are present at all altitudes, the microwave flux that reaches the satellite is an integral of emissions from thick layers of the atmosphere[i]. The observed microwave radiance, or "brightness temperature", is related to the average temperature of a broad layer of the atmosphere by a weighting function, which describes the relative contribution of each level of the atmosphere to the total radiance. The weighting function is calculated using an atmospheric radiative transfer model. The function depends both on the microwave frequency band that is observed and the angle of observation relative to Earth's surface, allowing the sounder to measure different layers in the atmosphere via the use of different frequency bands and/or different viewing angles (Mears and Wentz 2016; Zou et al. 2006; Po-Chedley et al. 2015).

We used satellite estimates of atmospheric temperature change produced by four different research groups:

1. Remote Sensing Systems in Santa Rosa, California (RSS; Mears et al. 2011;

[i]Satellite estimates of the temperature of tropospheric layers also receive a small contribution from the temperature at Earth's surface.

10.1175/JCLI-D-16-0333.s1

Mears and Wentz 2016).

2. The Center for Satellite Applications and Research, NOAA/National Environmental Satellite, Data, and Information Service, Camp Springs, Maryland (STAR; Zou et al. 2006, 2009; Zou and Wang 2011).

3. The University of Alabama at Huntsville (UAH; Christy et al. 2007).

4. The University of Washington, Seattle, Washington (UW; Po-Chedley et al. 2015).

All four groups provide satellite estimates of layer-average changes in the temperature of the mid- to upper troposphere (TMT). Three groups (RSS, UAH, and STAR) produce satellite measurements of the temperature of the lower stratosphere (TLS). Currently, only two groups (RSS and UAH) supply measurements of the temperature of the lower troposphere (TLT). The approximate altitude ranges and pressure level boundaries for TMT, TLS, and TLT are given in Table 2 in Karl et al. (2006).

The construction of a long-term data record from satellite data requires that measurements from more than a dozen satellites are intercalibrated and merged together. The four groups use different methods to account for: 1) the effects of inter-satellite differences in the MSU calibration; 2) calibration drift as the physical temperature of the satellite varies; 3) changes in the design and measurement frequencies of the MSU instrument itself; and 4) the effects of drifts in the time of day that the mea-

10.1175/JCLI-D-16-0333.s1

surements are made. In turn, these different analysis methods and processing choices yield substantial and important differences in the long-term trends reported by each group.

UAH provides two different versions (5.6 and 6.0 beta5) of their TLS, TMT, and TLT datasets. RSS currently has only one version (3.3) of their TLS and TLT datasets, but two versions (3.3 and 4.0) of their TMT product. Two versions were available for the STAR TLS and TMT datasets (3.0 and 4.0). There is currently only one version (1.0) of the UW tropical TMT dataset.

All satellite datasets were in the form of monthly means on 2.5° × 2.5° latitude/longitude grids. At the time this analysis was performed, satellite temperature data were available for the 450-month period from January 1979 to June 2016. For most calculations, we analyzed complete years only. For computing the trends in Fig. 6 and the trends and $p_c(i,k,l)'$ values in Fig. 7, we also included the first six months of 2016. Extension of our trend significance analysis to June 2016 illustrates the sharp decrease in $p_c(i,k,l)'$ values as the L-year sliding windows begin to sample the large tropospheric warming in early 2016.

There are differences in the spatial coverage of the satellite temperature data produced by the four groups.

1. UW TMT data are available for the tropics only (30°N-30°S).

10.1175/JCLI-D-16-0333.s1

2. All UAH satellite datasets have global coverage.

3. STAR TLS and TMT products extend from 87.5°N to 87.5°S.

4. RSS datasets extend from 82.5°N to 82.5°S for TLS and TMT, and from 82.5°N to 70°S for TLT.

There are two reasons why the RSS TLT coverage is restricted to 82.5°N to 70°S. First, there are virtually no temperature measurements poleward of 82.5° from the central view angle of the satellite "swath". Second, reliable estimation of brightness temperatures in the Southern Hemisphere is hampered by the large, poorly-known surface emissivity contribution from snow- and ice-covered areas of the Antarctic continent that are above 3,000 meters.

To exclude any impact of spatial coverage differences on trend comparisons, we calculated all near-global averages of actual and 'synthetic' satellite temperatures over the areas of common coverage in the RSS, UAH, and STAR datasets (82.5°N to 82.5°S for TLS and TMT, and 82.5°N to 70°S for TLT). All tropical averages are over 20°N to 20°S.

1.1 Calculating corrected TMT data from satellite datasets

The approach used to correct TMT data for stratospheric cooling is described in Appendix A. It relies on both TLS and TMT data (Fu et al. 2004; Fu and Johanson

10.1175/JCLI-D-16-0333.s1

2004). In calculating corrected TMT (referred to here as TMT_{cr}) from UAH TLS and TMT data, we did not 'mix' different versions of the UAH datasets: version 5.6 of UAH TMT_{cr} was computed with version 5.6 of UAH TLS and TMT data, and version 6.0 of UAH TMT_{cr} was computed with version 6.0 of UAH TLS and TMT.

For RSS, version 3.3 of TMT_{cr} was calculated with version 3.3 of RSS TLS and TMT data. Version 4.0 of RSS TMT_{cr} relied on version 4.0 of RSS TMT and version 3.3 of RSS TLS (since version 4.0 of RSS TLS is not yet available). The residual errors that were corrected in the transition from version 3.3 to version 4.0 of the RSS TMT data are unlikely to have pronounced impact on TLS, so the inconsistency in the TMT and TLS versions used to generate version 4.0 of the RSS TMT_{cr} data is not important (Mears and Wentz 2016).

The UW research group provides a tropical TMT dataset, but not a tropical TLS product. We were therefore unable to estimate TMT_{cr} using UW data only, and relied on TLS estimates produced by other groups[ii] to correct UW tropical TMT data. This yields five different versions of UW tropical TMT_{cr} data. We do not specifically identify these five individual dataset versions in Figs. 4 and 8.

[ii] RSS version 3.3, UAH versions 5.6 and 6.0, and STAR versions 3.0 and 4.0.

10.1175/JCLI-D-16-0333.s1

1.2 Satellite-average results

1.2.1 Calculation of satellite averages

We display "satellite-average" temperature time series[iii] in Fig. 1, and also discuss "satellite-average" values of the model/data trend ratios in Figs. 2, 4, and 5. In computing these averages, our strategy is to avoid giving undue weight to a particular research group with multiple dataset versions. For each research group with multiple dataset versions, we first average over individual versions. These "group averages" are then used in forming the overall average from different observational research groups. This is similar to the strategy we employ in calculation of multi-model averages, where we average over realizations before averaging over models.

For tropical TMT_{cr}, for example, there are a total of 11 individual dataset versions from 4 research groups (2 from RSS, 2 from UAH, 2 from STAR, and 5 from UW; see above). For near-global averages of TMT_{cr}, there are only 5 individual dataset versions from 3 research groups[iv]. Both tropical and near-global averages of TLS have 5 individual dataset versions from 3 research groups (RSS, UAH, and STAR). The smallest number of dataset versions (3) is for tropical and near-global averages

[iii]For the y-axis ranges used in Fig. 1, the observational temperature time series from individual research groups are virtually superimposed on each other. This is why we prefer to show the average observational temperature changes in Fig. 1.

[iv]This is because the UW TMT data are available for a tropical domain only.

10.1175/JCLI-D-16-0333.s1

of TLT, which are available from RSS and UAH only.

As mentioned in the main text and in Appendix B, we calculate two forms of the "satellite-average" trend ratio, $\overline{\overline{R}}$. The first relies on all available dataset versions, while the second is based on the most recent versions of each research group's TMT and TMT_{cr} data (see *Supplemental Material*, Table 5). The latter values of $\overline{\overline{R}}$ are consistently smaller and closer to unity, because more recent dataset versions show larger global-scale warming of the mid- to upper troposphere.

1.2.2 Stratification of satellite averages by volcanic forcing

Not all of the CMIP-5 models analyzed here incorporated volcanic forcing in their simulations of historical climate change (see *Supplemental Material*, Table 2). In calculating satellite-average atmospheric temperature changes, we used simulations with and without explicit treatment of the radiative effects of stratospheric volcanic aerosols, which we refer to subsequently as "VOLC" and "NoVOLC" runs[v]. The motivation for this choice was that two of the primary claims we seek to test[vi] also relied on an analysis of both VOLC and NoVOLC simulations. Our selection of CMIP-5 models emulated the choice of CMIP-5 models used to support the claims

[v]Note that the p-values in Fig. 7 do not rely on information from the ALL+8.5 simulations, and are therefore unaffected by inter-model differences in the treatment of historical volcanic forcing.

[vi]We test the claims that the satellite-average trend ratio $\overline{\overline{R}} = 3$ for near-global averages of TMT data and $\overline{\overline{R}} = 4$ for tropical averages of TMT (Christy 2015).

10.1175/JCLI-D-16-0333.s1

made in Christy (2015).

In the real world, however, atmospheric temperature records show pronounced lower stratospheric warming and tropospheric cooling after major volcanic eruptions (Santer et al. 2013a,b). In order to facilitate model comparisons with observational atmospheric temperature records, it is therefore scientifically justifiable to exclude "NoVOLC" models from the calculation of $\overline{\overline{R}}$. To assess how stratification of the ALL+8.5 simulations by treatment of volcanic forcing affects our $\overline{\overline{R}}$ results, we recalculated $\overline{\overline{R}}$ using a subset of 28 VOLC models[vii].

Consider first the case of trends in near-global averages of uncorrected TMT. In the 37-model "VOLC+NoVOLC" analysis, the timescale- and satellite-average trend ratio is 2.37 if $\overline{\overline{R}}$ is based on all six satellite dataset versions, and 2.08 if $\overline{\overline{R}}$ is calculated with the most recent versions of the RSS, STAR, and UAH datasets (*Supplemental Material*, Table 5). If trend ratios are based solely on the 28 VOLC models, the corresponding $\overline{\overline{R}}$ values are 2.39 and 2.10 – i.e., excluding NoVOLC models has minimal impact on $\overline{\overline{R}}$.

The situation is similar for trends in near-global averages of corrected TMT. Here,

---

[vii]The excluded NoVOLC models were: `CMCC-CESM1`, `CMCC-CM`, `CMCC-CMS`, `FGOALS-g2`, `FIO-ESM`, `INM-CM4`, `IPSL-CM5a-LR`, `IPSL-CM5a-MR`, and `IPSL-CM5b-LR`. The information from Table 2 of the *Supplemental Material* suggests that all of these models have some form of historical volcanic forcing, but the forcing does not involve explicit interaction of stratospheric volcanic aerosols with incoming solar and outgoing longwave radiation.

10.1175/JCLI-D-16-0333.s1

the full 37 models yield $\overline{\overline{R}} = 1.86$ and $\overline{\overline{R}} = 1.71$ for the "all satellite dataset" and "most recent satellite dataset" calculations. The $\overline{\overline{R}}$ values for the 28 VOLC models are only 2-3% larger (1.91 and 1.75, respectively)[viii]. We conclude from this sensitivity analysis that the trend ratios reported in our paper, which are based on the 37-model "VOLC+NoVOLC" analysis, are not markedly affected by the inclusion of simulations lacking explicit volcanic forcing.

2 Details of model output

2.1 General information

We used model output from phase 5 of the Coupled Model Intercomparison Project (CMIP-5; see Taylor et al. 2012). A list of modeling groups participating in CMIP-5 is given at http://cmip-pcmdi.llnl.gov/cmip5/docs/CMIP5_modeling_groups.pdf. The model simulations analyzed here were contributed by 19 different research groups (see *Supplemental Material*, Table 1). Our focus was on three different types of simulation:

[viii]The fact that $\overline{\overline{R}}$ is consistently slightly larger in the 28-model case than in the 37-model case is probably partly related to the absence of short-term volcanically induced stratospheric warming signals in the NoVOLC simulations. However, other factors (such as inter-model differences in stratospheric ozone forcing; see Santer et al. 2013b), complicate the interpretation of differences between the VOLC and VOLC+NoVOLC $\overline{\overline{R}}$ results.

10.1175/JCLI-D-16-0333.s1

1. Simulations with estimated historical changes in human and natural external forcings (see *Supplemental Material*, Table 2).

2. Simulations with 21st century changes in greenhouse gases and anthropogenic aerosols prescribed according to the Representative Concentration Pathway 8.5 (RCP8.5), with radiative forcing of approximately 8.5 W/m^2 in 2100, eventually stabilizing at roughly 12 W/m^2.

3. Pre-industrial control runs with no changes in external influences on climate, which provide information on internal climate noise.

Most CMIP-5 historical simulations end in December 2005. RCP8.5 simulations were typically initiated from conditions of the climate system at the end of the historical run. To avoid truncating comparisons between modeled and observed atmospheric temperature trends in December 2005, we spliced together synthetic satellite temperatures from the historical simulations and the RCP8.5 runs. Splicing allows us to compare actual and synthetic temperature changes over the full 37-year length of the satellite record. We use the acronym “ALL+8.5” to identify these spliced simulations.

Details of the start dates, end dates, and lengths of the historical simulations and RCP8.5 runs are given in *Supplemental Material*, Table 3. Corresponding information for the pre-industrial control runs is supplied in *Supplemental Material*, Table 4. In total, we analyzed 49 individual ALL+8.5 realizations, performed with 37 different CMIP-5 models. We relied on pre-industrial control runs from 36 CMIP-5 models.

10.1175/JCLI-D-16-0333.s1

2.2 Calculation of synthetic satellite temperatures

In many previous comparisons of modeled and observed atmospheric temperature trends, a global-mean weighting function was convolved with the atmospheric temperature profiles at each model grid-point. There is a different global-mean weighting function for each atmospheric layer of interest (TLS, TMT, and TLT). Here, we use a local weighting function method developed at RSS. At each model grid-point, the simulated temperature profiles were convolved with local weighting functions. The local weights depend on the grid-point surface pressure, the surface type (land or ocean), and the selected layer-average temperature (TLS, TMT, or TLT). This method provides more accurate estimates of synthetic satellite temperatures, particularly over high elevation regions (Santer et al. 2013b).

2.3 Issues related to model selection

2.3.1 Splicing of historical and RCP8.5 simulations

Christy (2015) has claimed that CMIP-5 TMT trends over 1979 to 2015 are a factor of 3 to 4 larger than satellite TMT trends. To address the validity of this claim, we analyze simulations performed with most of the same models that were used by Christy (2015). There are two notable differences. Christy (2015) relied on simulation output from the `CNRM-CM5` and `BNU-ESM` models, while we do not. Both `CNRM-CM5` and

10.1175/JCLI-D-16-0333.s1

BNU-ESM have a discontinuity in radiative forcing between the end of the historical simulation and the beginning of the RCP8.5 integration. In CNRM-CM5, stratospheric volcanic aerosol is assumed to decay to background values by the end of the historical simulation. At the beginning of all CNRM-CM5 RCP8.5 integration, an estimate of 'historical average' stratospheric aerosol optical depth (over 850 to 1999 A.D.) is continuously applied (Santer et al. 2013b). A similar 'historical average' volcanic aerosol loading was used in the BNU-ESM RCP simulations (Ji Duoying, personal communication).

These discontinuities in volcanic forcing induce abrupt and sustained warming of the lower stratosphere at the transition between the end of the historical run and the beginning of RCP8.5. Lower stratospheric warming has noticeable impact on the synthetic TLS trends computed with CNRM-CM5 and BNU-ESM, and hence affects the estimated TMT and TMT_{cr} trends obtained using these two models. We do not believe, therefore, that it is justifiable to include either CNRM-CM5 or BNU-ESM in comparisons of simulated and observed TMT and TMT_{cr} trends[ix].

[ix]In earlier work, we included CNRM-CM5 results in model-data trend comparisons (Santer et al. 2013b). This was feasible because "historicalExt" simulations were available for CNRM-CM5. These simulations enabled extension of the CNRM-CM5 historical run without any discontinuity in volcanic forcing. Unfortunately, the CNRM-CM5 "historicalExt" simulations end in December 2012, so CNRM-CM5 atmospheric temperatures that are unaffected by forcing discontinuities can no longer be compared with satellite data over the full 1979 to 2015 analysis period used here.

10.1175/JCLI-D-16-0333.s1

2.3.2 Treatment of GISS-E2-H and GISS-E2-R models

In the `GISS-E2-H` and `GISS-E2-R` models, the same atmospheric GCM is coupled to different ocean models. For each of these two coupled models, ALL+8.5 simulation output was available for different model versions (`p1` and `p3` for `GISS-E2-H`, and (`p1`, `p2`, and `p3` for `GISS-E2-R`). For the purposes of calculating multi-model average (MMA) quantities, it was necessary to decide whether atmospheric temperatures from these individual model versions should be treated as different realizations of historical climate change performed with a similar physical model, or as results from different models of the climate system[x].

There are important differences between these model versions. Historical and future changes in aerosols and ozone are prescribed in `p1`, but are interactive in `p2` and `p3` (Eyring et al. 2013; Shindell et al. 2013). Version `p2` has "an a priori calculation of the aerosol indirect effect", while `p3` uses a parameterized scheme (Eyring et al. 2013). Such differences can have significant implications for the atmospheric temperature changes (and the temperature variability) simulated in these model versions. We therefore decided to treat `p1`, `p2`, and `p3` as separate models.

This decision does not only impact the calculation of the MMA – it also affects

[x]In other words, whether we were dealing with five separate models [`GISS-E2-H (p1)`, `GISS-E2-H (p3)`, `GISS-E2-R (p1)`, `GISS-E2-R (p2)`, and `GISS-E2-R (p3)`], or with two realizations of `GISS-E2-H` and three realizations of `GISS-E2-R`.

10.1175/JCLI-D-16-0333.s1

estimates of internal variability. For GISS-E2-H, synthetic MSU temperatures were available from the p1, p2, and p3 control runs. For GISS-E2-R, synthetic MSU temperatures were available from the p1 and p2 control runs only (see *Supplemental Material*, Table 4). In the p-value calculations shown in Figs. 7 and 8, we treated the GISS atmospheric temperatures as estimates of internal variability from five different models (instead of regarding them as three realizations of GISS-E2-H and two realizations of GISS-E2-R).

3 Additional discussion of results

3.1 Noise differences in Figure 1

The "satellite-average" temperature time series in Fig. 1 are noticeably noisier than the CMIP-5 multi-model averages. This difference in 'noisiness' has a simple explanation. In satellite observations, there is only one sequence of the random noise of internal climate variability. Averaging over the atmospheric temperatures produced by different research groups does not damp this noise. In contrast, internal climate noise is not correlated (except by chance) in the 49 model realizations of historical climate change. In the "model world", therefore, averaging over realizations and models reduces the size of random climate noise, more clearly revealing the underlying atmospheric temperature response to external influences.

10.1175/JCLI-D-16-0333.s1

3.2 Compensating errors in TLS multi-model averages

As noted in the discussion of Figs. 1A and B, it is fortuitous that the peak lower stratospheric warming after major eruptions is similar in amplitude in satellite observations and in the CMIP-5 multi-model average. This apparent agreement arises from compensating errors. Most CMIP-5 models with explicit consideration of the radiative impact of stratospheric volcanic aerosols overestimate the observed peak lower stratospheric warming after El Chichón and Pinatubo (Santer et al. 2013b). This bias is approximately compensated for by the lack of 20th century volcanic forcing in a small number of CMIP-5 simulations (see *Supplemental Material*, Table 2).

3.3 Similarities in modeled and observed timescale-dependence of trends

One curious feature of Figs. 2A and 2B is that both $\overline{b_o}(k,l)$ and $\overline{\overline{b_f}}(l)$ show a broad peak with a maximum at $L \approx 22$ years. This common peak is related to two factors: the relatively short length of the time series, and the length of L in relation to the phasing of the maximum tropospheric cooling caused by Pinatubo (which occurs in mid-1992, nearly 22 years before the end of the TMT time series). Twenty-two-year TMT trends which commence in mid-1992 sample both the 'rebound' from Pinatubo-induced cooling and the underlying anthropogenic warming trend.

10.1175/JCLI-D-16-0333.s1

Supplemental Material Table 1: CMIP-5 models used in this study.

	Model	**Country**	**Modeling center**
1	ACCESS1.0	Australia	Commonwealth Scientific and Industrial Research Organization and Bureau of Meteorology
2	ACCESS1.3	Australia	Commonwealth Scientific and Industrial Research Organization and Bureau of Meteorology
3	BCC-CSM1.1	China	Beijing Climate Center, China Meteorological Administration
4	BCC-CSM1.1(m)	China	Beijing Climate Center, China Meteorological Administration
5	CanESM2	Canada	Canadian Centre for Climate Modelling and Analysis
6	CCSM4	USA	National Center for Atmospheric Research
7	CESM1-BGC	USA	National Science Foundation, U.S. Dept. of Energy, National Center for Atmospheric Research
8	CESM1-CAM5	USA	National Science Foundation, U.S. Dept. of Energy, National Center for Atmospheric Research
9	CMCC-CESM	Italy	Centro Euro-Mediterraneo per I Cambiamenti Climatici
10	CMCC-CM	Italy	Centro Euro-Mediterraneo per I Cambiamenti Climatici
11	CMCC-CMS	Italy	Centro Euro-Mediterraneo per I Cambiamenti Climatici
12	CSIRO-Mk3.6.0	Australia	Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence
13	EC-EARTH	Various	EC-EARTH consortium
14	FGOALS-g2	China	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences; and CESS, Tsinghua University
15	FIO-ESM	China	The First Institute of Oceanography, SOA
16	GFDL-CM3	USA	NOAA Geophysical Fluid Dynamics Laboratory
17	GFDL-ESM2G	USA	NOAA Geophysical Fluid Dynamics Laboratory

10.1175/JCLI-D-16-0333.s1

Supplemental Material Table 1: CMIP-5 models used in this study (continued).

	Model	**Country**	**Modeling center**
18	GFDL-ESM2M	USA	NOAA Geophysical Fluid Dynamics Laboratory
19	GISS-E2-H (p1)	USA	NASA Goddard Institute for Space Studies
20	GISS-E2-H (p2)	USA	NASA Goddard Institute for Space Studies
21	GISS-E2-H (p3)	USA	NASA Goddard Institute for Space Studies
22	GISS-E2-R (p1)	USA	NASA Goddard Institute for Space Studies
23	GISS-E2-R (p2)	USA	NASA Goddard Institute for Space Studies
24	GISS-E2-R (p3)	USA	NASA Goddard Institute for Space Studies
25	HadGEM2-CC	UK	Met. Office Hadley Centre
26	HadGEM2-ES	UK	Met. Office Hadley Centre
27	INM-CM4	Russia	Institute for Numerical Mathematics
28	IPSL-CM5A-LR	France	Institut Pierre-Simon Laplace
29	IPSL-CM5A-MR	France	Institut Pierre-Simon Laplace
30	IPSL-CM5B-LR	France	Institut Pierre-Simon Laplace
31	MIROC5	Japan	Atmosphere and Ocean Research Institute (the University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology
32	MIROC-ESM-CHEM	Japan	As for MIROC5
33	MIROC-ESM	Japan	As for MIROC5
34	MPI-ESM-LR	Germany	Max Planck Institute for Meteorology

10.1175/JCLI-D-16-0333.s1

Supplemental Material Table 1: CMIP-5 models used in this study (continued).

	Model	**Country**	**Modeling center**
35	MPI-ESM-MR	Germany	Max Planck Institute for Meteorology
36	MRI-CGCM3	Japan	Meteorological Research Institute
37	NorESM1-M	Norway	Norwegian Climate Centre
38	NorESM1-ME	Norway	Norwegian Climate Centre

10.1175/JCLI-D-16-0333.s1

Supplemental Material Table 2: External forcings in the historical simulations performed with the 37 CMIP-5 models used in this study. Information was extracted from the metadata of the relevant NetCDF files.*

	Model	Forcing information from metadata
1	ACCESS1.0	GHG, Oz, SA, Sl, Vl, BC, OC§
2	ACCESS1.3	GHG, Oz, SA, Sl, Vl, BC, OC§
3	BCC-CSM1.1	Nat, Ant, GHG, SD, Oz, Sl, Vl, SS, Ds, BC, OC
4	BCC-CSM1.1(m)	Nat, Ant, GHG, SD, Oz, Sl, Vl, SS, Ds, BC, OC
5	CanESM2	GHG, Oz, SA, BC, OC, LU, Sl, Vl
6	CCSM4	Sl, GHG, Vl, SS, Ds, SD, BC, MD, OC, Oz, AA, LU
7	CESM1-BGC	Sl, GHG, Vl, SS, Ds, SD, BC, MD, OC, Oz, AA, LU
8	CESM1-CAM5	Sl, GHG, Vl, SS, Ds, SD, BC, MD, OC, Oz, AA, LU
9	CMCC-CESM	Nat, Ant, GHG, SA, Oz, Sl
10	CMCC-CM	Nat, Ant, GHG, SA, Oz, Sl
11	CMCC-CMS	Nat, Ant, GHG, SA, Oz, Sl
12	CSIRO-Mk3.6.0	Ant, Nat
13	EC-EARTH	Nat, Ant
14	FGOALS-g2	GHG, Oz, SA, BC, Ds, OC, SS, Sl, Vl#
15	FIO-ESM	Nat, Ant
16	GFDL-CM3	GHG, SA, Oz, LU, Sl, Vl, SS, BC, MD, OC†
17	GFDL-ESM2G	GHG, SD, Oz, LU, Sl, Vl, SS, BC, MD, OC‡
18	GFDL-ESM2M	GHG, SD, Oz, LU, Sl, Vl, SS, BC, MD, OC‡
19	GISS-E2-H (p1)	GHG, LU, Sl, Vl, BC, OC, SA, Oz**
20	GISS-E2-H (p3)	GHG, LU, Sl, Vl, BC, OC, SA, Oz**
21	GISS-E2-R (p1)	GHG, LU, Sl, Vl, BC, OC, SA, Oz**
22	GISS-E2-R (p2)	GHG, LU, Sl, Vl, BC, OC, SA, Oz**
23	GISS-E2-R (p3)	GHG, LU, Sl, Vl, BC, OC, SA, Oz**
24	HadGEM2-CC	GHG, Oz, SA, LU, Sl, Vl, BC, OC

10.1175/JCLI-D-16-0333.s1

Supplemental Material Table 2 (continued): Historical external forcings.

Model		Forcing information from metadata
25	HadGEM2-ES	GHG, SA, Oz, LU, Sl, Vl, BC, OC[¶]
26	INM-CM4	GHG, Oz, SI, SA, Vl
27	IPSL-CM5A-LR	Nat, Ant, GHG, SA, Oz, LU, SS, Ds, BC, MD, OC, AA
28	IPSL-CM5A-MR	Nat, Ant, GHG, SA, Oz, LU, SS, Ds, BC, MD, OC, AA
29	IPSL-CM5B-LR	Nat, Ant, GHG, SA, Oz, LU, SS, Ds, BC, MD, OC, AA
30	MIROC5	GHG, SA, Oz, LU, Sl, Vl, SS, Ds, BC, MD, OC[††]
31	MIROC-ESM-CHEM	GHG, SA, Oz, LU, Sl, Vl, MD, BC, OC
32	MIROC-ESM	GHG, SA, Oz, LU, Sl, Vl, MD, BC, OC
33	MPI-ESM-LR	GHG, Oz, SD, Sl, Vl, LU
34	MPI-ESM-MR	GHG, Oz, SD, Sl, Vl, LU
35	MRI-CGCM3	GHG, SA, Oz, LU, Sl, Vl, BC, OC[‖]
36	NorESM1-M	GHG, SA, Oz, Sl, Vl, BC, OC
37	NorESM1-ME	GHG, SA, Oz, Sl, Vl, BC, OC

[*]Forcing abbreviations are described in Appendix 1.2 of the CMIP-5 Data Reference Syntax document. Nat = natural forcing (a combination, not explicitly defined); Ant = anthropogenic forcing (a combination, not explicitly defined); GHG = well-mixed greenhouse gases; SD = anthropogenic sulfate aerosol (direct effects only); SI = anthropogenic sulfate aerosol (indirect effects only); SA = anthropogenic sulfate aerosol direct and indirect effects; Oz = tropospheric and stratospheric ozone; LU = land-use change; Sl = solar irradiance; Vl = volcanic aerosol; SS = sea salt; Ds = dust; BC = black carbon; MD = mineral dust; OC = organic carbon; AA = anthropogenic aerosols (a mixture of aerosols, not explicitly defined).

[§]GHG = CO_2, N_2O, CH_4, CFC11, CFC12, CFC113, HCFC22, HFC125, HFC134a.

[#]GHG includes CO_2, CH_4, N_2O, CFC11, effective CFC12. Aerosol also includes sulfate.

[†]GHG includes CO_2, CH_4, N_2O, CFC11, CFC12, HCFC22, and CFC113. Aerosol direct and indirect effects are included.

[‡]GHG includes CO_2, CH_4, N_2O, CFC11, CFC12, HCFC22, and CFC113. "The direct effect of tropospheric aerosols is calculated by the model, but not the indirect effects."

[**]Also includes orbital change, BC on snow, and nitrate aerosols.

[¶]GHG = CO_2, N_2O, CH_4, CFCs.

[††]GHG includes CO_2, N_2O, CH_4, and CFCs; Oz includes OH and H_2O_2; LU excludes change in lake fraction.

[‖]GHG includes CO_2, CH_4, N_2O, CFC-11, CFC-12, and HCFC-22.

10.1175/JCLI-D-16-0333.s1

Supplemental Material Table 3: Basic information relating to the start dates, end dates, and lengths (N_m, in months) of the 49 CMIP-5 historical and RCP8.5 simulations used in this study. EM is the "ensemble member" identifier*.

	Model	EM	Hist. Start	Hist. End	Hist. N_m	RCP8.5 Start	RCP8.5 End	RCP8.5 N_m
1	ACCESS1.0	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
2	ACCESS1.3	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
3	BCC-CSM1.1	r1i1p1	1850-01	2012-12	1956	2006-01	2300-12	3540
4	BCC-CSM1.1(m)	r1i1p1	1850-01	2012-12	1956	2006-01	2099-12	1128
5	CanESM2	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
6	CanESM2	r2i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
7	CanESM2	r3i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
8	CanESM2	r4i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
9	CanESM2	r5i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
10	CCSM4	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
11	CCSM4	r2i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
12	CCSM4	r3i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
13	CESM1-BGC	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
14	CESM1-CAM5	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
15	CMCC-CESM	r1i1p1	1850-01	2005-12	1872	2000-01	2095-12	1140
16	CMCC-CM	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
17	CMCC-CMS	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
18	CSIRO-Mk3.6.0	r10i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
19	EC-EARTH	r8i1p1	1850-01	2012-12	1956	2006-01	2100-12	1140
20	FGOALS-g2	r1i1p1	1850-01	2005-12	1872	2006-01	2101-12	1152
21	FIO-ESM	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
22	GFDL-CM3	r1i1p1	1860-01	2005-12	1752	2006-01	2100-12	1140
23	GFDL-ESM2G	r1i1p1	1861-01	2005-12	1740	2006-01	2100-12	1140
24	GFDL-ESM2M	r1i1p1	1861-01	2005-12	1740	2006-01	2100-12	1140

10.1175/JCLI-D-16-0333.s1

Supplemental Material Table 3 (continued): Information on the 49 CMIP-5 historical and RCP8.5 simulations used in this study.

Model		EM	Hist. Start	Hist. End	Hist. (months)	RCP8.5 Start	RCP8.5 End	RCP8.5 (months)
25	GISS-E2-H (p1)	r1i1p1	1850-01	2005-12	1872	2006-01	2300-12	3540
26	GISS-E2-H (p3)	r1i1p3	1850-01	2005-12	1872	2006-01	2300-12	3540
27	GISS-E2-R (p1)	r1i1p1	1850-01	2005-12	1872	2006-01	2300-12	3540
28	GISS-E2-R (p2)	r1i1p2	1850-01	2005-12	1872	2006-01	2300-12	3540
29	GISS-E2-R (p3)	r1i1p3	1850-01	2005-12	1872	2006-01	2300-12	3540
30	HadGEM2-CC	r1i1p1	1859-12	2005-11	1752	2005-12	2099-12	1129
31	HadGEM2-CC	r2i1p1	1959-12	2005-12	553	2005-12	2099-12	1129
32	HadGEM2-CC	r3i1p1	1959-12	2005-12	553	2005-12	2099-12	1129
33	HadGEM2-ES	r1i1p1	1859-12	2005-11	1752	2005-12	2299-12	3529
34	HadGEM2-ES	r2i1p1	1859-12	2005-12	1753	2005-12	2100-11	1140
35	INM-CM4	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
36	IPSL-CM5A-LR	r1i1p1	1850-01	2005-12	1872	2006-01	2300-12	3540
37	IPSL-CM5A-LR	r2i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
38	IPSL-CM5A-MR	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
39	IPSL-CM5B-LR	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
40	MIROC5	r1i1p1	1850-01	2012-12	1956	2006-01	2100-12	1140
41	MIROC-ESM-CHEM	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
42	MIROC-ESM	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
43	MPI-ESM-LR	r1i1p1	1850-01	2005-12	1872	2006-01	2300-12	3540
44	MPI-ESM-LR	r2i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
45	MPI-ESM-LR	r3i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
46	MPI-ESM-MR	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
47	MRI-CGCM3	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
48	NorESM1-M	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140
49	NorESM1-ME	r1i1p1	1850-01	2005-12	1872	2006-01	2100-12	1140

*See http://cmip-pcmdi.llnl.gov/cmip5/documents.html for further details.

10.1175/JCLI-D-16-0333.s1

Supplemental Material Table 4: Start dates, end dates, and lengths (N_m, in months) of the 36 CMIP-5 pre-industrial control runs used in this study. EM is the "ensemble member" identifier.*

	Model	EM	Start	End	N_m
1	ACCESS1.0	r1i1p1	300-01	799-12	6000
2	ACCESS1.3	r1i1p1	250-01	749-12	6000
3	BCC-CSM1.1	r1i1p1	1-01	500-12	6000
4	BCC-CSM1.1(m)	r1i1p1	1-01	400-12	4800
5	CanESM2	r1i1p1	2015-01	3010-12	11952
6	CCSM4	r1i1p1	800-01	1300-12	6012
7	CESM-BGC	r1i1p1	101-01	600-12	6000
8	CESM-CAM5	r1i1p1	1-01	319-12	3828
9	CMCC-CESM	r1i1p1	4324-01	4600-12	3324
10	CMCC-CM	r1i1p1	1550-01	1879-12	3960
11	CMCC-CMS	r1i1p1	3684-01	4183-12	6000
12	CSIRO-Mk3.6.0	r1i1p1	1651-01	2150-12	6000
13	FGOALS-g2	r1i1p1	201-01	900-12	8400
14	FIO-ESM	r1i1p1	401-01	1200-12	9600
15	GFDL-CM3	r1i1p1	1-01	500-12	6000
16	GFDL-ESM2G	r1i1p1	1-01	500-12	6000
17	GFDL-ESM2M	r1i1p1	1-01	500-12	6000
18	GISS-E2-H (p1)	r1i1p1	2410-01	2949-12	6480
19	GISS-E2-H (p2)	r1i1p2	2490-01	3020-12	6372
20	GISS-E2-H (p3)	r1i1p3	2490-01	3020-12	6372
21	GISS-E2-R (p1)	r1i1p1	3981-01	4530-12	6600
22	GISS-E2-R (p2)	r1i1p2	3590-01	4120-12	6372
23	HadGEM2-CC	r1i1p1	1859-12	2099-12	2881
24	HadGEM2-ES	r1i1p1	1859-12	2435-11	6912
25	INM-CM4	r1i1p1	1850-01	2349-12	6000
26	IPSL-CM5A-LR	r1i1p1	1800-01	2799-12	12000

10.1175/JCLI-D-16-0333.s1

Supplemental Material Table 4 (continued): Information on the 36 CMIP-5 pre-industrial control runs used in this study.

	Model	EM	Start	End	N_m
27	IPSL-CM5A-MR[§]	r1i1p1	1800-01	2068-12	3228
28	IPSL-CM5B-LR	r1i1p1	1830-01	2129-12	3600
29	MIROC5	r1i1p1	2000-01	2669-12	8040
30	MIROC-ESM-CHEM	r1i1p1	1846-01	2100-12	3060
31	MIROC-ESM	r1i1p1	1800-01	2330-12	6372
32	MPI-ESM-LR	r1i1p1	1850-01	2849-12	12000
33	MPI-ESM-MR	r1i1p1	1850-01	2849-12	12000
34	MRI-CGCM3	r1i1p1	1851-01	2350-12	6000
35	NorESM1-M	r1i1p1	700-01	1200-12	6012
36	NorESM1-ME	r1i1p1	901-01	1152-12	3024

[*] See http://cmip-pcmdi.llnl.gov/cmip5/documents.html for further details.

[§] The `IPSL-CM5A-MR` control run has a large discontinuity in year 2069. We therefore truncated the `IPSL-CM5A-MR` control run after December 2068.

10.1175/JCLI-D-16-0333.s1

Supplemental Material Table 5: Values of the timescale-average trend ratio, $\overline{R}(k)$, from Figs. 2C,D and 4C,D of the main text. The final two rows show $\overline{\overline{R}}$, the trend ratios averaged over both satellite datasets and timescales. Results are for tropospheric temperatures averaged over a near-global domain (columns 2-4) and over the tropics (columns 5-6)¶.

Dataset	TMT (global)	TMT_{cr} (global; M_{cst})*	TMT_{cr} (global; M_{lat})§	TMT (tropical)	TMT_{cr} (tropical)
1 RSS v3.3	2.577	2.204	2.092	2.809	2.394
2 RSS v4.0	1.606	1.495	1.457	1.886	1.723
3 STAR v3.0	1.730	1.592	1.548	2.062	1.837
4 STAR v4.0	1.544	1.447	1.418	1.638	1.515
4 UAH v5.6	3.666	2.654	2.378	5.631	3.731
5 UAH v6.0	3.102	2.437	2.248	4.436	3.236
6 UW v1.0	–	–	–	2.236	≈ 1.955
OBS average (all datasets)	2.371	1.972	1.857	2.867	2.285
OBS average (new datasets)	2.084	1.793	1.708	2.549	2.107

¶For each of the k satellite datasets, $\overline{R}(k)$ is the ratio between the average of the distribution of externally forced model trends and the average of the distribution of satellite trends (averaged over all analysis timescales); $\overline{\overline{R}}$ is the overall average (over timescales and satellite datasets) of the individual $\overline{R}(k)$ values. Ratios based on TMT_{cr} data were corrected for the influence of stratospheric cooling. For further information, refer to Appendix B.

*The M_{cst} method of correcting TMT for stratospheric cooling relies on latitudinally invariant regression coefficients, with $a_{24} = 1.1$ (see Appendix A).

§The M_{lat} method of correcting TMT for stratospheric cooling relies on latitudinally varying regression coefficients, with $a_{24} = 1.1$ between 30°N and 30°S, and $a_{24} = 1.2$ poleward of 30° (see Appendix A).