THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Celebrating the Anniversary of Three Key Events in Climate Change Science

Citation for published version:
Santer, BD, Bonfils, CJW, Fu, Q, Fyfe, JC, Hegerl, G, Mears, C, Painter, JF, Po-Chedley, S, Wentz, FJ, Zelinka, MD & Zou, C-Z 2019, 'Celebrating the Anniversary of Three Key Events in Climate Change Science', Nature Climate Change. https://doi.org/10.1038/s41558-019-0424-x

Digital Object Identifier (DOI):
10.1038/s41558-019-0424-x

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Nature Climate Change

Celebrating the Anniversary of Three Key Events in Climate Change Science

Benjamin D. Santer[1], C.J.W. Bonfils[1], Qiang Fu[2], John C. Fyfe[3], Gabriele C. Hegerl[4], Carl Mears[5], Jeffrey F. Painter[1], Stephen Po-Chedley[1], Frank J. Wentz[5], Mark D. Zelinka[1] & Cheng-Zhi Zou[6]

[1]Program for Climate Model Diagnosis and Intercomparison (PCMDI), Lawrence Livermore National Laboratory, Livermore, CA 94550, USA.

[2]Dept. of Atmospheric Sciences, University of Washington, Seattle, WA 98195, USA.

[3]Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria, British Columbia, V8W 2Y2, Canada.

[4]School of Geosciences, University of Edinburgh, Edinburgh EH9 3FE, UK.

[5]Remote Sensing Systems, Santa Rosa, CA 95401, USA.

[6]Center for Satellite Applications and Research, NOAA/NESDIS, College Park, MD 20740, USA.

Corresponding author's email: santer1@llnl.gov

Submitted to *Nature Climate Change*

Date: January 1, 2019

Climate science celebrates three 40th anniversaries in 2019: release of the Charney report, publication of a key paper on anthropogenic signal detection, and the start of satellite temperature measurements. This confluence of scientific understanding and data led to the identification of a human fingerprint in atmospheric temperature.

We discuss below the events commemorated by these anniversaries. Our focus is on understanding how the scientific advances arising from these events aided efforts to identify human influences on the thermal structure of the atmosphere.

The Charney report

In 1979, the U.S. National Academy of Sciences published the findings of an "Ad Hoc Study Group on Carbon Dioxide and Climate". This is frequently referred to as the Charney report[1]. The authors did not have many of the scientific advantages available today: international climate science assessments based on thousands of relevant peer-reviewed scientific papers[2,3,4], four decades of satellite measurements of global climate change[5], land and ocean surface temperature datasets spanning more than 120 years[6], estimates of natural climate variability[7,8], and sophisticated three-dimensional numerical models of Earth's climate system. Nevertheless, the report's principal findings have aged remarkably well. Consider conclusions regarding the equilibrium climate sensitivity (ECS): "*We estimate the most probable global warming for a doubling of CO_2 to be near $3°C$ with a probable error of +/- $1.5°C$*". These values are in accord with current understanding[9] and are now supported by multiple lines of evidence that were unavailable in 1979. Examples include observed patterns of surface warming, greenhouse gas and temperature changes on Ice Age timescales, and results from multi-model ensembles of externally forced simulations[3,4,9].

There is also better process-level understanding of the feedbacks contributing to ECS uncertainties[10,11,12]. Charney *et al.* understood that the factor of three spread in ECS was mainly due to uncertainties in the net effect of high and low cloud feedbacks[13]. Reliable assessment of cloud feedbacks required "*comprehensive numerical modeling of the general circulations of the atmosphere and the oceans together with validation by comparison of the observed with the model-produced cloud types and amounts.*" This conclusion foreshadowed rigorous evaluation of model cloud properties with satellite data[14]. Such comparisons ultimately led to the elucidation of robust cloud responses to greenhouse warming[15], and to the 2013 conclusion of the Intergovernmental Panel on Climate Change (IPCC) that "*the sign of the net radiative feedback due to all cloud types is... likely positive*"[10].

The ocean's role in climate change featured prominently in the Charney report. The authors noted that ocean heat uptake would delay the emergence of a human-caused warming signal from the background noise of natural variability[16]. This delay, they wrote, meant that humanity "*...may not be given a warning until the CO_2 loading is such that an appreciable climate change is inevitable*". The finding that "*On time scales of decades... the coupling between the mixed layer and the upper thermocline must be considered*" provided impetus for the development of atmosphere-ocean General Circulation Models (GCMs).

The authors also knew that scientific uncertainties did not negate the reality and seriousness of human-caused climate change: "*We have examined with care all known negative feedback mechanisms, such as increase in low or middle cloud amount, and have concluded that the oversimplifications and inaccuracies in the models are not likely to have vitiated the principal conclusion that there will be appreciable warming.*" Although the GCMs available in 1979 were not yet sufficiently reliable for predicting regional changes, Charney *et al.* cautioned that the "*associated regional climate changes so important to the assessment of socioeconomic consequences may well be significant*".

In retrospect, the Charney report seems like the scientific equivalent of the handwriting on the wall. Forty years ago, its authors issued a clear warning of the potentially significant socioeconomic consequences of human-caused warming. Their warning was accurate, and remains more relevant than ever.

Hasselmann's optimal detection paper

The second scientific anniversary marks the publication of a paper by Klaus Hasselmann entitled: "*On the signal-to-noise problem in atmospheric response studies*"[17]. This is now widely regarded as the first serious effort to provide a sound statistical framework for identifying a human-caused warming signal.

In the 1970s, there was recognition that GCM simulations yielded both "signal" and "noise" when forced by changes in atmospheric CO_2 or other external factors[18]. The signal was the climate response to the altered external factor. The noise arose from natural internal climate variability. Noise estimates were obtained from observations or by running an atmospheric GCM coupled to a simple model of the upper ocean. In the presence of intrinsic noise, statistical methods were required to identify areas of the world where first detection of a human-caused warming signal might occur.

One key insight in Hasselmann's 1979 paper was that analysts should look at the statistical significance of global geographical patterns of climate change. Previous work had assessed the significance of the local climate response to a particular external forcing at thousands of individual model grid-points. Climate information at these individual locations was correlated in space and in time, hampering assessment of overall significance. Hasselmann noted that "*...it is necessary to regard the signal and noise fields as multi-dimensional vector quantities... and the significance analysis should accordingly be carried out with respect to this multi-variate statistical field, rather than in terms of individual gridpoint statistics*". Instead of looking for the needle in a tiny corner of a large haystack (and then proceeding to search the next tiny corner), Hasselmann advocated for a more efficient strategy – searching the entire haystack simultaneously.

The paper pointed out that theory, observations, and models provide considerable information about signal and noise properties. For example, changes in solar irradiance, volcanic aerosols, and greenhouse gases produce signals with different patterns, amplitudes, and frequencies[2,3,4,8,19]. These unique signal characteristics ("fingerprints") can be used to distinguish climate signals from climate noise.

Hasselmann's paper was a statistical roadmap for hundreds of subsequent climate change detection and attribution ("D&A") studies. These investigations identified anthropogenic fingerprints in a wide range of independently monitored observational datasets[2,3,4]. D&A research provided strong scientific support for the conclusion reached by the IPCC in 2013: "*it is extremely likely that human influence has been the dominant cause of the observed warming since the mid-20th century*"[4].

Forty years of satellite temperature data

In November 1978, Microwave Sounding Units (MSUs) on NOAA polar-orbiting satellites began monitoring the microwave emissions from oxygen molecules. These emissions are proportional to the temperature of broad atmospheric layers[5]. A successor to MSU, the Advanced Microwave Sounding Unit (AMSU), was deployed in 1998. Estimates of global changes in atmospheric temperature can be obtained from MSU and AMSU measurements.

Over their 40-year history, MSU and AMSU data have been essential ingredients in hundreds of research investigations. These datasets allowed scientists to study the size, significance, and causes of global trends and variability in Earth's atmospheric temperature and circulation, to quantify the tropospheric cooling after major volcanic eruptions, to evaluate climate model performance, and to assess the consistency between observed surface and tropospheric temperature changes[2,3,4,20].

Satellite atmospheric temperature data were also a useful test-bed for Hasselmann's signal detection strategy. They had continuous, near-global coverage[5]. Data products were available from multiple research groups, allowing uncertainties in temperature retrievals to be quantified. Signal detection studies with MSU and AMSU revealed that human fingerprints were identifiable in the warming of the troposphere and cooling of the lower stratosphere[8], confirming model projections made over 50 years ago[21]. Tropospheric warming is largely due to increases in atmospheric CO_2 from fossil fuel use[2,3,4,8,20], while lower stratospheric cooling over the 40-year satellite record[22] is mainly attributable to anthropogenic depletion of stratospheric ozone[23].

While enabling significant scientific advances, MSU and AMSU temperature data have also been at the center of scientific and political imbroglios. Some controversies were related to differences between surface warming inferred from thermometers and tropospheric warming estimated from satellites. Claims that these warming rate differences cast doubt on the reliability of the surface data have not been substantiated[20,24]. Other disputes focused on how to adjust for non-climatic artifacts arising from orbital decay and drift, instrument calibration drift, and the transition between MSU and AMSU instruments[5,20]. More recently, claims of no significant warming since 1998 have been based on artfully selected subsets of satellite temperature data. Such claims are erroneous and do not call into question the reality of long-term tropospheric warming[25].

A confluence of scientific understanding

The zeitgeist of 1979 was favorable for anthropogenic signal detection. From the Charney report, which relied on basic theory and early climate model simulations, there was clear recognition that fossil fuel burning would yield an appreciable global warming signal[1]. Klaus Hasselmann's paper[17] outlined a rational approach for detecting this signal. Satellite-borne microwave sounders began to monitor atmospheric temperature, providing global patterns of multi-decadal climate change and natural internal variability – information required for successful application of Hasselmann's signal detection method.

Because of this confluence in scientific understanding, we can now answer the following question: when did a human-caused tropospheric warming signal first emerge from the background noise of natural climate variability? We addressed this question by applying a fingerprint method related to Hasselmann's approach (see online Methods). An anthropogenic fingerprint of tropospheric warming is identifiable with high statistical confidence in all currently available satellite datasets (Figure 1). In two out of three datasets, fingerprint detection at a 5-sigma threshold – the gold standard for discoveries in particle physics – occurs no later than 2005, only 27 years after the 1979 start of the satellite measurements. Humanity cannot afford to ignore such clear signals.

References

1. Charney, J. G. *et al.* Carbon dioxide and climate: A scientific assessment. Available from Climate Research Board, National Research Council (1979).

2. Mitchell, J. F. B. & Karoly, D. J. Detection of climate change and attribution of causes. In J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, editors, Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, pages 695–738. Cambridge University Press (2001).

3. Hegerl, G. C. *et al.* Understanding and Attributing Climate Change. In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, editors, Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Inter-governmental Panel on Climate Change, pages 663–745. Cambridge University Press (2007).

4. Bindoff, N. L. *et al.* Detection and Attribution of Climate Change: from Global to Regional. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, editors, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pages 867–952. Cambridge University Press, (2013).

5. Mears, C. & Wentz, F. J. *J. Clim.* **30**, 7695–7718 (2017).

6. Morice, C. P., Kennedy, J. J., Rayner, N. A. & Jones, P. D. *J. Geophys. Res.* **117**, (2012).

7. Fyfe, J. C. *et al.*. *Nat. Comm.* **8**, (2017).

8. Santer, B. D. *et al. Proc. Nat. Acad. Sci.* **110**, 17235–17240 (2013).

9. Knutti, R., Rugenstein, M. A. A., & Hegerl, G. C. *Nat. Geosci.* **10**, 727-736 (2017).

10. IPCC. Summary for Policymakers. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, editors, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press (2013).

11. Ceppi, P., Brient, F., Zelinka, M. D., & Hartmann, D. L. *WIREs Clim. Change* **8**, (2017).

12. Caldwell, P. M., Zelinka, M. D., Taylor, K. E., & Marvel, K. *J. Clim.* **29**, 513–524 (2016).

13. Klein, S. A., Hall, A., Norris, J. R., & Pincus, R. *Surv. Geophys.* **38**, 1307–1329 (2017).

14. Klein, S. A., Zhang, Y., Zelinka, M. D., Pincus, R., Boyle, J., & Gleckler, P. J. *J. Geophys. Res.* **118**, 1329–1342 (2013).

15. Zelinka, M. D., Randall, D. A., Webb, M. J., & Klein, S. A. *Nat. Clim. Change* **7**, 674–678 (2017).

16. Barnett, T. P. *et al. Science* **309**, 284–287 (2005).

17. Hasselmann, K.. On the signal-to-noise problem in atmospheric response studies. In Meteorology over the Tropical Oceans, pages 251–259. Roy. Met. Soc., London, 1979.

18. Chervin, R. M., Washington, W. M., & Schneider, S. H. *J. Atmos. Sci.* **33**, 413–423 (1976).

19. North, G. R., Kim, K. Y., Shen, S. S. P., & Hardin, J. W. *J. Clim.* **8**, 401-408 (1995).

20. Karl, T. R., Hassol, S. J., Miller, C. D., & Murray, W. L., eds. Temperature trends in the lower atmosphere: Steps for understanding and reconciling differences. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research. National Oceanic and Atmospheric Administration, 164 pages (2006).

21. Manabe, S. & Wetherald, R. T. *J. Atmos. Sci.* **24**, 241–259 (1967).

22. Zou, C.-Z. & Qian, H. *J. Atmos. Ocean. Tech.* **33**, 1967–1984 (2016).

23. Solomon, S. *et al. J. Geophys. Res.* **122**, 8940–8950 (2017).

24. Fu, Q., Johanson, C. M., Warren, S. G., & Seidel, D. J. *Nature* **429**, 55–58 (2004).

25. Santer, B. D. *et al. Sci. Reports* **7**, (2017).

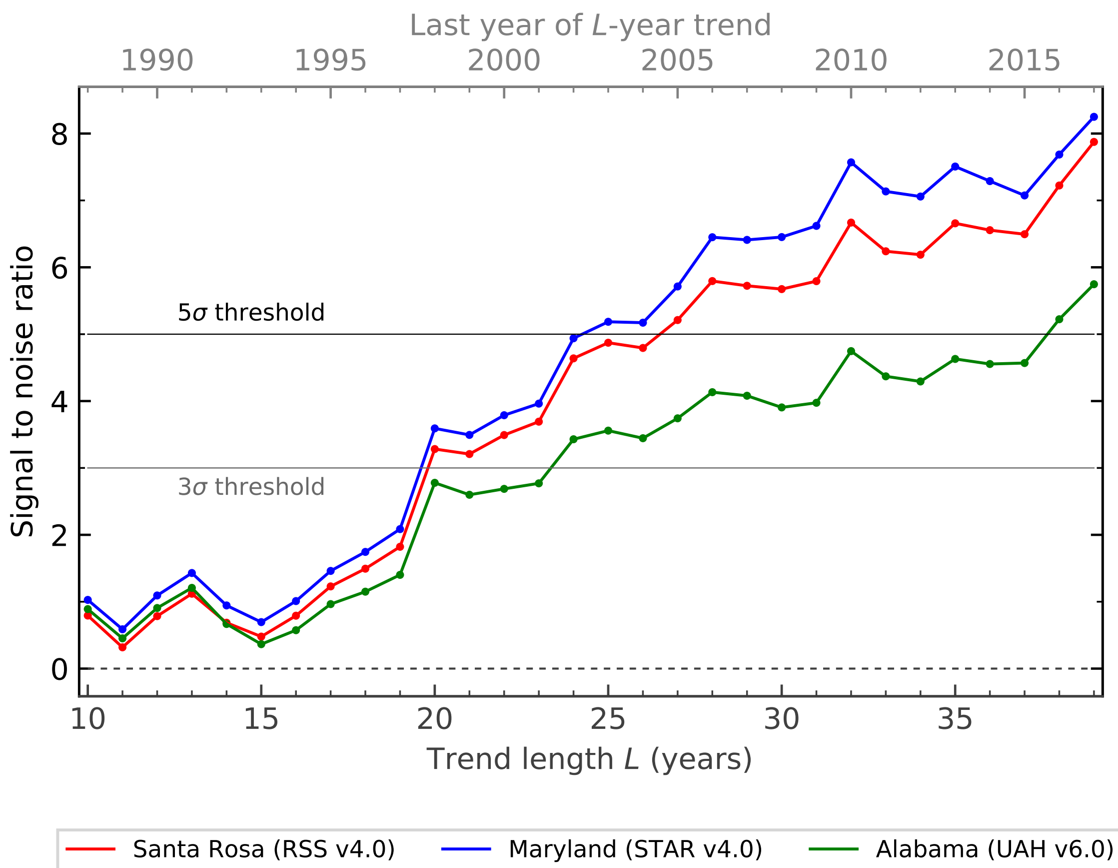

Figure 1: Signal-to-noise ratios (S/N) used for identifying a model-predicted anthropogenic fingerprint in satellite measurements of annual-mean tropospheric temperature. The MSU and AMSU measurements are from three different research groups: Remote Sensing Systems (RSS), the Center for Satellite Applications and Research (STAR), and the University of Alabama at Huntsville (UAH). The grey and black horizontal lines are the 3σ and 5σ thresholds that we use for estimating the signal detection time. By 2002, all three satellite datasets yield S/N ratios exceeding the 3σ threshold. By 2016, an anthropogenic signal is consistently detected at the 5σ threshold. Further details of the model and satellite data and the fingerprint method are provided in the online Methods.

Acknowledgments

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison (PCMDI) provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. The authors thank Susan Solomon (M.I.T.) and Ken Denman, Norm McFarlane, and Knut von Salzen (Canadian Centre for Climate Modelling and Analysis) for helpful comments. **Funding:** Work at LLNL was performed under the auspices of the U.S. Department of Energy under contract DE-AC52-07NA27344 through the Regional and Global Model Analysis Program (B.D.S., J.F.P., S.P.-C., and M.Z.) and the Early Career Research Program Award SCW1295 (C.B.). Support was also provided by NASA Grant NNH12CF05C (F.J.W. and C.M.), and NOAA Grant NA18OAR4310423 (Q.F). **Author contributions:** B.D.S. conceived the study and performed statistical analyses. J.F.P. calculated synthetic satellite temperatures from model simulation output. C.M., F.J.W., and C.-Z.Z. provided satellite temperature data. All authors contributed to the writing and revision of the manuscript. **Competing interests:** None. **Data and materials availability:** All primary satellite and model temperature datasets used here are publicly available. Derived products (synthetic satellite temperatures calculated from model simulations) are provided at: https://pcmdi.llnl.gov/research/DandA/. **Disclaimer:** The views, opinions, and findings contained in this report are those of the authors and should not be construed as a position, policy, or decision of the U.S. Government, the U.S. Department of Energy, or the National Oceanic and Atmospheric Administration.

Online Methods

1 Satellite atmospheric temperature data

In calculating the signal detection times shown in Figure 1, we used satellite estimates of atmospheric temperature produced by Remote Sensing Systems[5,26], the Center for Satellite Applications and Research[27,28], and the University of Alabama at Huntsville[29,30]. We refer to these groups subsequently as RSS, STAR, and UAH (respectively). All three groups provide satellite measurements of the temperatures of the mid- to upper troposphere (TMT) and the lower stratosphere (TLS). Our focus here is on estimating the detection time for an anthropogenic fingerprint in satellite TMT data. TLS is required for correcting TMT for the influence it receives from stratospheric cooling[24] (see Section 3).

Satellite datasets are in the form of monthly means on $2.5^\circ \times 2.5^\circ$ latitude/longitude grids. At the time this analysis was performed, temperature data were available for the 468-month period from January 1979 to December 2017. We used the most recent dataset versions from each group: 4.0 (RSS), 4.0 (STAR), and 6.0 (UAH).

We note that studies of the size, patterns, and causes of atmospheric temperature changes have also relied on information from radiosondes[20,31,32,33,34]. Non-climatic factors, such as refinements over time in radiosonde instrumentation and thermal shielding, hamper the identification of true climate changes[20,35]. Additionally, radiosonde data have much sparser coverage than satellite data, particularly in the Southern Hemisphere. The spatially complete coverage of MSU and AMSU offers advantages for obtaining reliable estimates of hemispheric- and global-scale temperature trends and patterns of temperature change.

2 Details of model output

We used model output from phase 5 of CMIP, the Coupled Model Intercomparison Project[36]. The simulations analyzed here were contributed by 19 different research groups (see Supplementary Table S1). Our focus was on three different types of numerical experiment: 1) simulations with estimated historical changes in human and natural external forcings; 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; and 3) pre-industrial control runs with no changes in external influences on climate. Details of these simulations are provided in Supplementary Tables S2 and S3.

Most CMIP5 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 39-year length of the satellite record. We use the acronym "HIST+8.5" to identify these spliced simulations.

3 Method used for correcting TMT data

Trends in TMT estimated from microwave sounders receive a substantial contribution from the cooling of the lower stratosphere[24,37,38,39]. In Fu et al. (2004)[24], a regression-based method was developed for removing the bulk of this stratospheric cooling component of TMT. This method has been validated with both observed and model atmospheric temperature data[37,40,41]. Here, we refer to the corrected version of TMT as TMT_{cr}. The main text discusses corrected TMT only, and does not use the subscript cr to identify corrected TMT.

For calculating tropical averages of TMT_{cr}, Fu et al. (2005)[38] used:

$$\text{TMT}_{cr} = a_{24}\text{TMT} + (1 - a_{24})\text{TLS} \tag{1}$$

where $a_{24} = 1.1$. For the global domain considered here, lower stratospheric cooling makes a larger contribution to TMT trends, so a_{24} is larger[24,39]. In Fu et al (2004)[24] and Johanson and Fu (2006)[39], $a_{24} \approx 1.15$ was applied directly to near-global averages of TMT and TLS. Since we are performing corrections on local (grid-point) data, we used $a_{24} = 1.1$ between $30°$N and $30°$S, and $a_{24} = 1.2$ poleward of $30°$. This is approximately equivalent to use of the $a_{24} = 1.15$ for globally-averaged data.

4 Calculation of synthetic satellite temperatures

We use a local weighting function method developed at RSS to calculate synthetic satellite temperatures from model output[42]. At each model grid-point, simulated temperature profiles were convolved with local weighting functions. The weights depend on the grid-point surface pressure, the surface type (land or ocean), and the selected layer-average temperature (TLS or TMT).

5 Fingerprint method

Detection methods generally require an estimate of the true but unknown climate-change signal in response to an individual forcing or set of forcings[16,17,43,44,45,46]. This is often referred to as the fingerprint $F(x)$.

We define $F(x)$ as follows. Let $S(i, j, x, t)$ represent annual-mean synthetic MSU temperature data at grid-point x and year t from the i^{th} realization of the j^{th} model's HIST+8.5 simulation, where:

$i = 1, \ldots N_r(j)$ (the number of realizations for the j^{th} model).

$j = 1, \ldots N_m$ (the number of models used in fingerprint estimation).

$x = 1, \ldots N_x$ (the total number of grid-points).

$t = 1, \ldots N_t$ (the time in years).

Here, N_r ranges from 1 to 5 realizations and $N_m = 37$ models. After transforming synthetic MSU temperature data from each model's native grid to a common $10^\circ \times 10^\circ$ latitude/longitude grid, $N_x = 576$ grid-points for corrected TMT. N_t is 39 years. We note that because the RSS TMT data do not have coverage poleward of 82.5°, the latitudinal extent of the regridded data is from $80°$N to $80°$S. This is the minimum common coverage in the three satellite datasets.

The multi-model average atmospheric temperature change, $\overline{\overline{S}}(x, t)$, was calculated by first averaging over an individual model's HIST+8.5 realizations (where multiple realizations were available), and then averaging over models. The double overbar denotes these two averaging steps. Anomalies were then defined at each grid-point x and year t with respect to the local climatological annual mean. The fingerprint $F(x)$ is the first Empirical Orthogonal Function (EOF) of the anomalies of $\overline{\overline{S}}(x, t)$. $F(x)$ was estimated over 1979 to 2017, the same time period used for determining observed TMT changes.

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

We obtain these variability estimates from control runs performed with multiple models. Because the length of the 36 control runs analyzed here varies by a factor of up to 4, models with longer control integrations could have a disproportionately large impact on our noise estimates. To guard against this possibility, the noise estimates rely on the last 200 years of each model's pre-industrial control run, yielding 7,200 years of concatenated control run data. Use of the last 200 years reduces the contribution of any initial residual drift to noise estimates.

Synthetic TMT data from individual model control runs are regridded to the same $10^\circ \times 10^\circ$ target grid used for fingerprint estimation. After regridding, anomalies are defined relative to the local climatological annual means calculated over the full length of each control run. Since control run drift can bias S/N estimates, its removal is advisable. Here, we assume that drift behavior can be well-approximated by a least-squares linear trend, and drift is removed at each grid-point. Drift removal is performed over the last 200 control run years (since only the last 200 years are concatenated).

Observed annual-mean TMT data are first transformed to the $10^\circ \times 10^\circ$ latitude/longitude grid used for the model HIST+8.5 simulations and control runs, and are then expressed as anomalies relative to climatological annual means over 1979 to 2017. The observed temperature data are projected onto $F(x)$, the time-invariant fingerprint:

$$Z_o(t) = \sum_{x=1}^{N_x} O(x,t)\, F(x) \quad t = 1, 2, \ldots, 39 \tag{2}$$

where $O(x,t)$ denotes the observed annual-mean TMT data. This projection is equivalent to a spatially uncentered covariance between the patterns $O(x,t)$ and $F(x)$ at year t. The signal time series $Z_o(t)$ provides information on the fingerprint strength in the observations. If observed patterns of temperature change are becoming increasingly similar to $F(x)$, $Z_o(t)$ should increase over time. A recent publication[47] provides figures showing both $F(x)$ and the observed patterns of annual-mean trends in TMT (see Figure S5A and Figures

2A,C, and E in Santer *et al.*, 2018).

Hasselmann's 1979 paper discusses the rotation of $F(x)$ in a direction that maximizes the signal strength relative to the control run noise[17]. Optimization of $F(x)$ generally leads to enhanced detectability of the signal[48,49]. In all cases we considered, we achieved detection of an externally-forced fingerprint in satellite TMT data without any optimization of $F(x)$. We therefore show only non-optimized results in our Figure 1.

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[50].

6 Estimating detection time

We assess the significance of changes in $Z_o(t)$ by comparing trends in $Z_o(t)$ with a null distribution of trends. To generate this null distribution, we require a case in which $O(x,t)$ is replaced by a record in which we know *a priori* that there is no expression of the fingerprint, except by chance. Here, we replace $O(x,t)$ by the concatenated noise data set $C(x,t)$, after first regridding and removing residual drift from $C(x,t)$ (as described above). The noise time series $N_c(t)$ is the projection of $C(x,t)$ onto the fingerprint:

$$N_c(t) = \sum_{x=1}^{N_x} C(x,t)\, F(x) \quad t = 1, \ldots, 7200 \tag{3}$$

Our detection time T_d is based on the signal-to-noise ratio, S/N. As in our previous work[47], we calculate S/N ratios by fitting least-squares linear trends of increasing length L years to $Z_o(t)$, and then comparing these with the standard deviation of the distribution of non-overlapping L-length trends in $N_c(t)$. Thus the numerator of the S/N ratio measures the trend in the pattern agreement between the model-predicted "human influence" fingerprint and observations; the denominator measures the trend in agreement between the fingerprint and patterns of natural climate variability. Detection occurs after L_d years, when the S/N ratio first exceeds some stipulated signal detection threshold, and then remains continuously above that threshold for all values of $L > L_d$. For example, $L_d = 10$ would signify that $T_d = 1988$ – *i.e.*, that detection of a human-caused tropospheric warming fingerprint occurred in 1988, 10 years after the start of the satellite temperature record.

We estimated T_d with both 3σ and 5σ signal detection thresholds. A 3σ threshold was used by Hansen et al. (1988) for detection of an anthropogenic signal in surface temperature[51]. A more stringent 5σ threshold is often employed as the gold standard for scientific discovery in particle physics* For detection at a 3σ threshold, there is a chance of roughly one in 741 that the "match" between the model-predicted anthropogenic fingerprint and the observed patterns of tropospheric temperature change could actually be due to natural internal variability (as represented by the 36 models analyzed here). With a 5σ detection threshold, this probability decreases to roughly one in 3.5 million†.

*For example, in detecting the existence of the Higgs boson. See, e.g., https: //understandinguncertainty.org/explaining-5-sigma-higgs-how-well-did-they-do.

†These are so-called complementary cumulative probabilities – see, e.g., https: //en.wikipedia.org/wiki/

We make three assumptions in order to calculate T_d. First, we assume that our knowledge of observed tropospheric temperature change is derived from the latest MSU and AMSU dataset versions produced by RSS, UAH, and STAR. Second, we assume that large ensembles of forced and unforced simulations performed with state-of-the-art climate models provide the best current estimates of a human fingerprint and natural internal climate variability[47]. Third, we assume that although the strength of the fingerprint in the observations changes over time, the fingerprint pattern itself is relatively stable[47].

At the 3σ threshold, $T_d = 1998$ for RSS and STAR and 2002 for UAH (Figure 1). This means that L_d is 20 years for RSS and STAR and 24 years for UAH. With a more stringent 5σ threshold the detection time is longer: $T_d = 2003$ for STAR, 2005 for RSS, and 2016 for UAH, yielding L_d values of 25, 27, and 38 years, respectively. The UAH results are noteworthy. Even though UAH tropospheric temperature data have consistently shown less warming than other datasets[24,52,53,54], UAH still yields confident 5σ detection of an anthropogenic fingerprint.

Finally, we note that detection times for an anthropogenic signal in surface temperature are available elsewhere[43,44,45,51] and have been the topic of recent discussion[‡].

Standard_normal_table#Cumulative. Probabilities are based on a one-tailed test. A one-tailed test is appropriate here, since we seek to determine whether natural variability could yield larger time-increasing similarity with the fingerprint pattern than the similarity we obtained by comparing the fingerprint with satellite data.

[‡]Another scientific anniversary received considerable attention in 2018 – the 30th anniversary of the publication of of a seminal paper by Jim Hansen and his colleagues at the NASA/Goddard Institute for Space Studies[51]. Some of the recent reporting on the 1988 Hansen et al. paper focused on the paper's prediction that "the global greenhouse warming should rise above the level of natural climate variability within the next several years". This prediction was for global-mean changes in surface temperature. It relied on a comparison of observed changes with multiple estimates of natural variability.

References for Online Methods

26. Mears, C. & Wentz, F. J. *J. Clim.* **29**, 3629–3646 (2016).

27. Zou, C.-Z. Zou & Wang, W. *J. Geophys. Res.* **116**, (2011).

28. Zou, C.-Z., Goldberg, M. D., & Hao, X. *Sci. Adv.* **4**, (2018).

29. Christy, J. R., Norris, W. B., Spencer, R. W., & Hnilo, J. J. *J. Geophys. Res.* **112**, (2007).

30. Spencer, R. W., Christy, J. R., & Braswell, W. D. *Asia-Pac. J. Atmos. Sci.* **53**, 121–130 (2017).

31. Thorne, P. W. *et al. Geophys. Res. Lett.* **29** (2002).

32. Thorne, P.W., Lanzante, J. R., Peterson, T. C., Seidel, D. J., & Shine, K. P. *Wiley Inter. Rev.* **2**, 66–88 (2011).

33. Lott, F. C. *et al.*. *J. Geophys. Res. Atmos.* **118**, 2609–2619 (2013).

34. Seidel, D. J. *et al. J. Geophys. Res.* **121**, 664–681 (2016).

35. Sherwood, S. C., Lanzante, J. R., & Meyer, C. L. *Science* **309**,155–1559 (2005).

36. Taylor, K. E., Stouffer, R. J., & Meehl, G. A. *Bull. Amer. Meteor. Soc.* **93**, 485–498 (2012).

37. Fu, Q. & Johanson, C. M. *J. Clim.* **17**, 4636–4640 (2004).

38. Fu, Q. & Johanson, C. M. *Geophys. Res. Lett.* **32** (2005).

39. Johanson, C. M. & Fu, Q. *J. Clim.* **19**, 4234–4242 (2006).

40. Gillett, N. P., Santer, B. D., & Weaver, A. J. *Nature* **432** (2004).

41. Kiehl, J.T., Caron, J., & Hack, J. J. *J. Clim.* **18**, 2533–2539 (2005).

42. Santer, B. D. *et al. Proc. Nat. Acad. Sci.* **110**, 26–33 (2013).

43. Hegerl, G. C. *et al. J. Clim.* **9**, 2281–2306 (1996).

44. Tett, S. F. B., Stott, P. A., Allen, M. R., Ingram, W. J., & Mitchell, J. F. B. *Nature* **399**, 569–572 (1999).

45. Stott, P. A. *et al. Science* **290**, 2133–2137 (2000).

46. Gillett, N. P., Zwiers, F. W., Weaver, A. J., & Stott, P. A. *Nature* **422**, 292-294 (2003).

47. Santer, B. D. *et al.*. *Science* **361** (2018).

48. Allen, M. R. & Tett, S. F. B. *Cli. Dyn.* **15**, 419–434 (1999).

49. Santer, B. D. *et al. Science* **300**, 1280–1284 (2003).

50. van den Dool, H.M., Saha, S., & Johansson, Å. *J. Clim.* **13**, 1421–1435 (2000).

51. Hansen, J. *et al. J. Geophys. Res.* **93**, 9341–9364 (1988).

52. Wentz, F. J. & Schabel, M. *Nature* **394**, 661–664 (1998).

53. Mears, C. A. & Wentz, F. J. *Science* **309**, 1548–1551 (2005).

54. Po-Chedley, S., Thorsen, T. J., & Fu, Q. *J. Clim.* **28**, 2274–2290 (2015).