

2012

# Statistical evidence about human influence on the climate system

---

Francisco Estrada, Pierre Perron, Benjamin Martinez. "Statistical evidence about human influence on the climate system."

<https://hdl.handle.net/2144/26280>

*"Downloaded from OpenBU. Boston University's institutional repository."*

**Statistically-derived contributions of diverse human influences to 20<sup>th</sup> century temperature changes**

Francisco Estrada<sup>1,2</sup>, Pierre Perron<sup>3</sup> & Benjamín Martínez-López<sup>1</sup>

*<sup>1</sup>Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Ciudad Universitaria, Circuito Exterior, 04510 Mexico, DF, Mexico, <sup>2</sup>Institute for Environmental Studies, Vrije Universiteit, Amsterdam, Netherlands, <sup>3</sup>Department of Economics, Boston University, 270 Bay State Rd. Boston, MA, 02215, USA.*

**The warming of the climate system is unequivocal as illustrated by an increase of global temperatures near 0.8°C during the last century. However, the attribution of the observed warming to human activities has remained debatable, particularly after the apparent slowdown since the 1990s. Applying recent statistical methods for trending time series with slope changes to radiative forcing and temperature series, we show that while other radiative factors modulated their effect, greenhouse gases defined the secular movement in total radiative forcing and temperatures. The salient feature is a marked increase in the growth rates of both temperatures and radiative forcing occurring near 1960, marking the start of sustained global warming. The trend imparted by radiative forcings also reveals the contribution of human interventions for two slowdowns in the warming. The Montreal Protocol and a reduction in methane emissions contributed to the recent one, while the two World Wars and the Great Crash contributed to the mid-20th century cooling via important reductions in CO<sub>2</sub> emissions. Results provide evidence about the effectiveness of reducing emissions of greenhouse gases not only to decrease long-term temperature change but also the rate of warming in the short-term.**

Two main statistical approaches are used to investigate the attribution of climate change: the optimal fingerprinting method<sup>1</sup> which consists in searching for spatial and/or temporal patterns consistent with the anthropogenic forcing signal that are common to observed and externally forced simulations of climate variables, and the cointegration framework that permits testing for the attribution of climate change directly from observed temperature and radiative forcing data<sup>2</sup>. As shown, temperature and radiative forcings are not integrated processes once breaks in trend are accounted for, rendering the latter approach inappropriate<sup>3,4</sup>. We therefore use recently developed statistical methods to analyze the properties of trending series<sup>5,6,7,8,9</sup>. We focus on providing evidence for the existence of a nonlinear trend characterized by breaks in slope that are common to observed global and hemispheric temperatures and anthropogenic forcing, thereby establishing direct evidence for the effect of human factors in altering the long-run path of global and hemispheric temperatures. Once this nonlinear trend is accounted for, all remaining variations in temperatures are stationary with different durations explained mostly by non-human factors. Our results are robust to different choices for temperatures and mixtures of anthropogenic and natural forcing series.

### **Analysis of the observed warming trends in temperatures**

The data for Global (G), Southern Hemisphere (SH) and Northern Hemisphere (NH) temperatures are from the NASA database and the Climatic Research Unit HadCRUT4, see Fig 1. Results for the HadCRUT3 dataset are also presented in the Supplementary Information. The forcing variables are the radiate forcing of greenhouse gases (RFGHG) mainly produced by anthropogenic activities, TRF\* which contains all

radiative forcing variables except the two main natural sources, solar and stratospheric aerosols, and TRF that includes TRF\* plus solar forcing and represents a mixture of all natural and anthropogenic radiative forcing variables that are trending (see Methods).

The Atlantic Multidecadal Oscillation (AMO) represents ocean-atmosphere processes naturally occurring in the North Atlantic with a large influence over NH and G climates<sup>10,11</sup>. It produces 60- to 90-years natural oscillations that distort the warming trend suggesting it should be filtered before conducting attribution studies<sup>10,12</sup>. After detrending G and NH with TRF, the residuals are highly correlated with AMO, indicating the importance of this mode of variability for explaining the low frequency variability in G and NH (Supplementary Information S1). Consequently, we remove the low frequency natural component of the AMO in order to obtain a better measure of the low frequency trend, i.e., to isolate the trend in climate. The filtered G and NH and unfiltered SH temperature series are graphed in Fig 1. A visual inspection clearly suggests nonlinear trend functions with an abrupt change in the warming rates<sup>3,13</sup>.

To statistically document the presence of a break in the trend of temperature series, we use the Perron-Yabu<sup>7</sup> testing procedure, valid with integrated or stationary noise, circumventing the problem of pre-testing for unit roots. The null hypothesis of no-break is rejected in all cases at the 1% significance level. Consider the regression

$$y_t = \mu + \beta_1 t + \beta_2 DT_t^* + \tilde{y}_t \quad (1)$$

where  $DT_t^* = t - T_B$  if  $t > T_B$  and zero otherwise,  $T_B$  denotes the time of the break,  $t$  is a time trend,  $\beta_2$  is the change in slope and  $\tilde{y}_t$  is the noise component. Minimizing the sum of squared residuals of regression (1), the break dates for G, and NH from the HadCRUT4 dataset are estimated to occur in 1956 and 1966 but for SH there is evidence of two breaks, in 1909 and 1976, the second being much larger. The NASA

dataset leads to similar estimates: 1956 for G, 1968 for NH and for SH in 1923 and 1955. Hence, all series show clear evidence of a break associated with a large increase in their growth rate around 1960 (Supplementary Information S2.4).

The data generating processes can be investigated using unit root tests that allow for a one-time break in the trend function. The Kim-Perron test<sup>6,5</sup> provides strong evidence that all temperature series are trend stationary processes accounting for the documented break, in accordance with results reported for observed and simulated temperatures<sup>3,14</sup> and those in S2 of the Supplementary Information, except that once the AMO is filtered the estimates of the break dates are not statistically different<sup>15</sup>. An exception is SH from HadCRUT4 for which the estimate of the break date is 1976. As discussed in S1 of the Supplementary Information, the difference between SH from the HadCRUT3 and HadCRUT4 appears to be characterized by an AMO-like low frequency oscillation. Interestingly, if the effect of AMO is filtered, the estimate of the break date for SH from HadCRUT4 is 1955 as for SH from NASA and HadCRUT3. In what follows, we shall continue using 1976 as the break date given that the results are robust to using 1955 or 1976.

All series show a slight warming until the mid-20th century: for the HadCRUT4 data, the increase is 0.30°C, 0.35°C, and 0.27°C per century for G, NH and SH, respectively; for NASA the corresponding figures are 0.31°C, 0.39°C, and 0.27°C. At the estimated break dates, the warming rates roughly tripled, with increases of 0.97°C, 1.18°C and 1.09°C per century for HadCRUT4 G, NH and SH, and corresponding increases of 0.94°C, 1.16°C, and 0.93°C for NASA. The estimates of the post-break

warming rates are roughly 1°C per century for G, NH, and SH for both datasets. After the break dates the warming has been uniform across hemispheres, as expected from an increase in radiative forcing of well-mixed greenhouse gases<sup>16</sup> once part of the thermal inertia and feedback effects that lead to strong warming differences between land and ocean have been removed by filtering the effects of AMO (Supplementary Information S1). Therefore, the large differences with the estimates of the warming rates and break dates reported using the unfiltered series are caused by the low-frequency natural variability mostly associated with AMO<sup>3,13,14</sup>.

### **Common trends in temperatures and radiative forcing**

We provide statistical evidence that the same features are present in the forcing variables, depicted in Fig 2. They also clearly show a nonstationary behavior with varying growth rates though, as expected, with much less short term variability<sup>17</sup>. Applying the same methodologies, the results indicate that all series are trend stationary processes with a highly significant break in growth rate estimated to occur in 1960 for RFGHG, TRF\* and TRF. This date is not statistically different from those of the slope breaks in the filtered temperature series (see Supplementary Table 10 for the confidence intervals). Hence, the temperature and forcing series have stationary noise components around trend functions with nearly common significant breaks in trend slope, indicating a secular co-movement.

To formally test for a common nonlinear trend in temperatures and radiative forcing a nonparametric nonlinear co-trending test<sup>9</sup> was applied to two sets composed of the G, NH, SH temperature series from each dataset, RFGHG and TRF. In both cases the results indicate the existence of four co-trending vectors involving G, NH, SH, RFGHG and TRF, hence a single common nonlinear trend. This suggests a dominant

anthropogenic influence on observed warming. The simplest radiative forcing series considered in each group is RFGHG which contains the nonlinear trend present in all other series, indicating it is the dominant driver imparting the common nonlinear trend to TRF and in turn to G, NH, and SH (Supplementary Table 11).

Fig 1 shows the fitted temperature series obtained from OLS regressions with different forcing series as the explanatory variable and with the filtered G and NH, and the unfiltered SH as the dependent variables. The concordance between the forcing and temperature trends is apparent; in all cases the fitted values appear the product of a low-pass filter on temperatures. Visual inspection of the residuals in Supplementary Figures 6a-6d strongly suggests that in all cases all secular movements have cancelled, leaving only stationary variations around zero.

The origin of the common trend can be established using the same testing procedure applied to the radiative forcing variables. The results show that the main secular movement of TRF\* and TRF is imparted by RFGHG, indicating that human-induced factors are the main drivers behind the observed warming.

The structural model behind this statistical models can be described by means of a simple two-compartments climate model<sup>18,19</sup>, for which due to small heat capacity and short time constant for reaching its steady state, the temperature in the atmosphere and the upper ocean follows closely the movements of external forcing. The slope coefficient relating these two variables is the transient climate sensitivity estimated to be  $0.35^{\circ}C(W/m^2)^{-1}$  and  $0.40^{\circ}C(W/m^2)^{-1}$  for NASA and HadCRUT4, respectively, concordant with previously reported values<sup>18</sup> (see Methods and Supplementary Information S6).

Our results have implications for previous work questioning the relevance of anthropogenic factors to explain the observed warming<sup>20</sup>. Their main conclusion is that previous cointegration-based studies<sup>2,21,29</sup> have overlooked the differences in the order of integration arguing that radiative forcings are integrated of order two while temperatures are integrated of order one. Given our findings, these conflicting results can easily be explained. If, as we claim, both series are trend-stationary with a change in slope, the application of standard unit root tests that misspecify the trend as linear will lead to their spurious results given the magnitude of the noise for each series<sup>5,22</sup> (see Supplementary Information S4, which also includes comparisons with other studies).

#### **Anthropogenic influence on the slowdowns in warming**

A relevant implication of our results is that deviations from the secular movement of temperatures are transitory: temperatures are reverting to the underlying trend determined by anthropogenic activities. Therefore, analyzing the radiative forcing trend provides a way to investigate smaller variations in the rate of warming that are obscured by the large natural temperature variability relative to the warming signal. The 1940-1970 cooling period and the recent slowdown in warming are of special interest.

Although describing temperature and forcing variables as having piecewise linear trends is convenient to investigate their time-series properties, it is a large simplification of the more complex secular movement. The analysis so far revealed that the overwhelming feature of the data is a large change in growth rates near 1960, variations are stationary around these broken trends and radiative forcings and temperatures have a common breaking trend; the break needs not be abrupt, a smooth transition is possible.

This does not preclude other nonlinearities not detected by statistical tests, such as a transition period or other small breaks.

The cooling period 1940-1970 has been explained as a mixture of natural variability (mainly AMO) and the cooling effect of anthropogenic aerosols produced during the industrial recuperation of Europe after WWII<sup>23,24,25,26,27,17</sup>. However, even when the effects of AMO are filtered, visual inspection still suggests a slowdown in warming but, importantly, with a shorter duration (1938-1955). Also, while the effect of aerosols is mainly limited to NH, the no-warming period also applies to G and SH<sup>25</sup> (about  $-0.07^{\circ}\text{C}$  per decade for G, NH and SH from HadCRUT4).

While these factors contributed to the cooling period, the  $\text{CO}_2$  radiative forcing (RFCO2) is another cause rarely mentioned. RFGHG from 1938 to early 1950's experienced a considerable slowdown in growth rate, remaining almost flat during 1938-1947, mainly from a decrease in RFCO2 for almost a decade (1940s, Supplementary Figure 7b). Otherwise, the cooling effect of anthropogenic aerosols (RAER) would have been mostly compensated by the increase in RFGHG: the yearly rate of growth of RAER during this cooling period was  $-0.013 \text{ W/m}^2$  while that of RFGHG before the slowdown in RFCO2 was  $0.010 \text{ W/m}^2$ .

The percent reduction in  $\text{CO}_2$  emissions during 1914-1946 has no parallel since 1751. These reductions were driven by three landmark socioeconomic events occurring in a brief period: the two World Wars and the largest economic crisis. Fig 3 depicts that during this period negative yearly growth rates were frequent in the gross domestic products of Europe and the USA and in global  $\text{CO}_2$  emissions. In particular, the Great Depression of 1929 induced a considerable reduction in world emissions of  $\text{CO}_2$  and other GHG related to economic activity. The largest drop in  $\text{CO}_2$  emissions occurred

between 1929 and 1932 (26% reduction) recovering its previous level only in 1937<sup>28</sup>. This led to a four-year period of negative changes in CO<sub>2</sub> and RFGHG (Fig 3), resulting in an inflection point around 1938 in RFGHG and TRF\* (Fig 2). The post-WWII economic expansion and the corresponding sharp and uninterrupted increase in anthropogenic forcing led to the occurrence of the common break in radiative forcing and temperature series around 1960, marking the beginning of sustained global warming.

The causes of a slowdown in warming since the mid-90s have been the object of interest. Some proposed the joint effect of increased short-lived sulfur emissions, La Niña events and the eleven-year solar cycle as offsetting the effect of rising greenhouse gases concentrations<sup>29</sup>. We show that the effects of the Montreal Protocol and of changes in agricultural practices in Asia have been large enough to change the long-run path of radiative forcing. Tropospheric aerosols contributed to making this slowdown more pronounced.

The causes of the reduced growth rate of RFGHG are twofold: the reduction of CFCs emissions and the "pause" in the growth rate of atmospheric methane<sup>30,31,32,13</sup>. The first is a direct consequence of the Montreal Protocol (1989) for controlling substances depleting the stratospheric ozone layer. The second is not completely understood but appears related to the decrease in microbial sources caused by the application of chemical fertilizers and to more efficient water use for producing rice in Asia<sup>32</sup>.

To obtain statistical evidence about the causes of the smaller increase in temperatures, we search for additional breaks in the post-1960 period using the sequential Perron-Yabu testing procedure developed by Kejriwal and Perron<sup>8,7</sup>. The slowdown in RFGHG is confirmed by a highly significant reduction of 25.61% in trend

slope in 1994. This break is even more pronounced when the direct and indirect effects of tropospheric aerosols are considered, as in TRF\* and TRF. The breaks are highly significant, occur in 1992 and 1991, respectively, and the rates of increase were reduced by more than 50% (the estimated break dates are not statistically different). Fig 2 presents the forcing series with the fitted trends obtained using the two estimated break dates. A similar break is also present in G from NASA: when considering the post 1960 sample, there is a statistically significant break in slope in the mid-1990's consistent with the reported slowdown in warming. This evidence is strong in view of the fact that the sample is quite short, though the result is not robust to using a longer sample. This change in the growth rate of radiative forcing is another common feature of the forcing and temperatures trends.

To obtain a better assessment, we applied the sequential break point detection procedure to all components of RFGHG, RAER, the indirect effects of aerosols (AIE) and black carbon (BC). Four components of RFGHG showed a statistically significant break in the post 1990 period (Fig 4). The largest change occurred for CFC<sub>11</sub>, with a break in 1993 and 116% decrease in slope, reverting to a sustained decrease. The second in magnitude occurred for CFC<sub>12</sub> with a break in 1995 and a 92.85% reduction in slope. These results provide clear evidence that the Montreal protocol was successful in achieving global reductions in CFC emissions. Although not its objective, the reductions were large enough to have an impact on RFGHG, which slowed the increase in warming<sup>31</sup>. The third largest decrease occurs for CH<sub>4</sub>, with a break in 1992 and the slope decreasing 73.35%. The last component with a break in the 1990s is CO<sub>2</sub> but it actually exhibits a 19.78% increase in slope in 1996. Hence, the evidence shows that the decrease in CFC<sub>11</sub>, and the reduced increase rate of CFC<sub>12</sub> and CH<sub>4</sub> are the major

contributors to the decrease in the growth rate of TRF, despite the more rapid increase in CO<sub>2</sub>. As discussed previously<sup>30,29</sup>, another significant factor for the slowdown in warming is the negative effect of the indirect effect of tropospheric aerosols which shows a 32.48% steeper slope since 1988 (Fig 4). Without these breaks in the components of the RFGHG, TRF would had been 0.25 W/m<sup>2</sup> larger in 2010 (about 0.13 W/m<sup>2</sup>, 0.05 W/m<sup>2</sup> and 0.08 W/m<sup>2</sup> for CH<sub>4</sub>, CFC<sub>11</sub> and CFC<sub>12</sub>, respectively), a small amount compared to the anthropogenic RFGHG but equivalent to a full amplitude solar forcing<sup>33</sup>, and about 15% of the increase in TRF since 1880 (Fig 5). If additionally the break in the indirect effect of tropospheric aerosols is removed, TRF would have been 0.34 W/m<sup>2</sup> larger, about a fifth of its increase from preindustrial times. Stratospheric aerosols from volcanic eruptions (e.g., Mount Pinatubo) cannot be responsible for a long lasting change as they have a short-lived effect on temperatures (see Supplementary Information S2.1)

Paradoxically, the recent decrease in warming, presented by global warming skeptics as proof that humankind cannot affect the climate system, is shown to have a direct human origin.

## **METHODS**

**Data.** The annual temperature data used are from the HadCRUT4 (1850-2010) (<http://www.metoffice.gov.uk/hadobs/hadcrut4/data/current/download.html>) and the GISS-NASA (1880-2010) datasets (<http://data.giss.nasa.gov/gistemp/>). The AMO (1856-2010) was obtained from NOAA (<http://www.esrl.noaa.gov/>). For the analysis in S1 the following climate indices were used: Southern Oscillation Index (SOI <http://www.cgd.ucar.edu/cas/catalog/climind/SOI.signal.ascii>); North Atlantic

Oscillation (NAO; <http://climexp.knmi.nl/data/inao.dat>); Pacific Decadal Oscillation (PDO; <http://jisao.washington.edu/pdo/PDO.latest>). For the sensitivity analysis in SI, the HadCRUT3 database was used (<http://www.metoffice.gov.uk/hadobs/hadcrut3/>).

We also used series from databases related to climate model simulations by the Goddard Institute for Space Studies (GISS-NASA). The radiative forcing data obtained from GISS-NASA (<http://data.giss.nasa.gov/modelforce/RadF.txt>) for the period 1880-2010 include the following (in  $\text{W/m}^2$ ): well-mixed greenhouse gases (carbon dioxide, methane, nitrous oxide and chlorofluorocarbons); ozone; stratospheric water vapor; solar irradiance; land use change; snow albedo; stratospheric aerosols; black carbon; reflective tropospheric aerosols; and the indirect effect of aerosols. The aggregated radiative forcing series were constructed as follows: RFGHG is the radiative forcing of the well-mixed greenhouse gases; TRF\* is RFGHG plus the radiative forcing of ozone, stratospheric water vapor, land use change; snow albedo, black carbon, reflective tropospheric aerosols and the indirect effect of aerosols; TRF is TRF\* plus solar irradiance. For the sensitivity analysis in S7, the direct effect of atmospheric aerosols was obtained from the RCP database (<http://www.pik-potsdam.de/~mmalte/rcps/index.htm#Download>). To analyze the individual components of RFGHG, the global mean mixing ratios of carbon dioxide, methane and chlorofluorocarbons were obtained from GISS-NASA (<http://data.giss.nasa.gov/modelforce/ghgases//>; [http://data.giss.nasa.gov/modelforce/ghgases//TG\\_A.1992-2009.txt](http://data.giss.nasa.gov/modelforce/ghgases//TG_A.1992-2009.txt)) and the radiative forcing due to these gases was calculated using simplified expressions (ref. 34). The global emissions from fossil-fuel burning, cement manufacture and gas flaring are from

the Carbon Dioxide Information Analysis Center ([http://cdiac.ornl.gov/trends/emis/tre\\_glob\\_2008.html](http://cdiac.ornl.gov/trends/emis/tre_glob_2008.html)). The GDP data (1880-2008) was obtained from the University of Groningen (<http://www.ggdc.nl/maddison/>).

**Structural and time series models.** The time-series models presented here have the general form:

$$T_t = \alpha + \gamma F_t + \varepsilon_t$$

where  $T_t$  is temperature,  $F_t$  is some measure of radiative forcing and  $\varepsilon_t$  is a climate noise encompassing both short-term and longer-term variability modes. The structural model can be described by a simple two-compartments climate model (ref. 18,19). The upper compartment is composed mainly of the atmosphere and the upper ocean and has a small heat capacity and short time constant to reach its steady state. This upper compartment is thermally coupled to the lower compartment, composed of the deep ocean, having a large heat capacity and long time constant for reaching the steady state. When a positive and sustained external forcing is applied to the system, the upper compartment temperature increases inducing changes in the absorbed and/or emitted radiation at the top of the atmosphere and a heat flow to the lower compartment, which has a much larger heat capacity and requires much longer time to respond to any forcing. The analyses presented in this paper relate to the response of the upper compartment of the climate system to increases in radiative forcing. The transient climate sensitivity characterized by the short time constant of the upper compartment is

$$S_{tr} = (\kappa + \lambda)^{-1}$$

where  $\kappa$  is the heat uptake coefficient of the climate system<sup>43</sup>. The transient climate sensitivity relates the time-dependent increase in surface temperature to the time-

dependent forcing such that  $\Delta T(t) = S_{ir}F(t)$  and it is equal to the slope parameter  $\gamma$  in the time-series model above. The response of the climate system to the forcing over the observed period is determined by the time constant of the upper compartment and the transient climate sensitivity, providing a physical explanation of why global and hemispheric surface temperatures follow the same nonlinear trend of the radiative forcing and why observed temperatures rapidly adjust to changes in the trend of the radiative forcing. These features are particularly clear after part of the low-frequency oscillations produced by AMO are filtered out of both G and NH. The overlapping confidence intervals in the break dates in the slope functions of radiative forcing and temperature series found in this paper are also consistent with the short time constant dominating this relationship, giving physical support to the idea of co-breaking in surface temperature and radiative forcing series. Supplementary Information S6 also presents a structural interpretation of the trend-stationary nature of aggregate radiative forcing.

## References

1. Hasselmann, K. Multi-pattern fingerprint method for detection and attribution of climate change. *Clim. Dyn.* **13**, 601– 611 (1997).
2. Kaufmann, R. K. & Stern, D.I. Evidence for human influence on climate from hemispheric temperature relations. *Nature* **388**, 39–44 (1997).
3. Gay, C., Estrada, F. & Sánchez, A. Global and hemispheric temperature revisited. *Clim Change* **94**, 333–349 (2009).

4. Estrada, F., Gay, C. & Sánchez, A. A reply to "Does temperature contain a stochastic trend? Evaluating conflicting statistical results". *Clim Change* **101**, 407–414 (2010).
5. Perron, P. The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* **57**, 1361–1401 (1989).
6. Kim, D. & Perron, P. Unit root tests allowing for a break in the trend function under both the null and the alternative hypotheses. *J. Econom.* **148**, 1-13 (2009).
7. Perron, P. & Yabu, T. Testing for shifts in trend with an integrated or stationary noise component. *JBES* **27**, 369-396 (2009).
8. Kejriwal, M. & Perron, P., 2009. A sequential procedure to determine the number of breaks in trend with an integrated or stationary noise component. *J. Time Ser. Anal.* **31**, 305–328 (2010).
9. Bierens, H. J. Nonparametric nonlinear cointegration analysis, with an application to interest and inflation in the United States. *JBES* **18**, 323-337 (2000).
10. Wu, Z., Huang, N. E., Wallace, J. M., Smoliak, B. V. & Chen, X. On the time-varying trend in global-mean surface temperature. *Clim. Dyn.* **37**, 759–773 (2011).
11. Swanson, K. L., Sugihara, G. & Tsonis, A. A. Long-term natural variability and the 20th century climate change. *Proc. Natl. Acad. Sci.* **106**, 16120-16123 (2009).
12. Knudsen, M. F., Seidenkrantz, M. S., Jacobsen, B. H. & Kuijpers, A. Tracking the Atlantic Multidecadal Oscillation through the last 8,000 years. *Nat. Comm.* **2**, 178 (2011).
13. IPCC: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt,

M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp (2007).

14. Estrada, F., Perron, P., Gay, C. & Martínez, B. A time series analysis of the 20th century climate simulations produced for the IPCC's AR4. *PLoS ONE* **8(3)**, e60017 (2013).

15. Perron, P. & Zhu, X. Structural breaks with deterministic and stochastic trends, *J. Econom.* **129**, 65-119 (2005).

16. Meehl, G. A., Washington, W. M., Wigley, T. M. L., Arblaster, J. M. & Dai, A. Solar and greenhouse gas forcing and climate response in the 20th century. *J. Climate* **16**, 426-444 (2003).

17. Hansen, J. & Sato, M. Greenhouse gas growth rates. *Proc. Natl. Acad. Sci.* **101**, 16109-16114 (2004).

18. Schwartz, S.E. Determination of Earth's transient and equilibrium climate sensitivities from observations over the twentieth century: strong dependence on assumed forcing. *Surv. Geophys.* **33**, 745–777 (2012).

19. Gregory, J.M. & Forster, P.M. Transient climate response estimated from radiative forcing and observed temperature change. *J. Geophys. Res.* **113**, D23105 (2008).

20. Beenstock, M., Reingewertz, Y., & Paldor, N.. Polynomial cointegration tests of anthropogenic impact on global warming. *ESD* **3**, 173-188 (2012).

21. Kaufmann, R.K., Kauppi, H. & Stock, J.H. Emissions, concentrations, & temperature: a time series analysis. *Clim. Change* **77**, 249-278 (2009).

22. Perron, P. Testing for a unit root in a time series with a changing mean. *JBES* **8**, 153-162 (1990).
23. Hansen, J. & Lebedeff, S. Global trends of measured surface air temperature. *J. Geophys. Res.* **92**, 13 345–13 372 (1987).
24. Jones, P.D., Raper, S.C.B., Bradley, R.S., Diaz, H.F., Kelly, P.M. & Wigley, T.M.L. Northern Hemisphere surface air temperature variations: 1851-1984. *J. Clim. App. Meteorol.* **25**, 161-179 (1986).
25. Jones, P.D., Raper, S.C.B. & Wigley, T.M.L. Southern Hemisphere surface air temperature variations: 1851-1984. *J. Clim. App. Meteorol.* **25**, 1213-1230 (1986).
26. Jones, P.D., Wigley, T.M.L. & Wright, P.B. Global temperature variations between 1861 and 1984. *Nature* **322**, 430-434 (1986).
27. Thompson, D.W.J., Kennedy, J.J., Wallace, J.M. & Jones, P.D. A large discontinuity in the mid-twentieth century in observed global mean surface temperature. *Nature* **453**, 646–649 (2008).
28. Andres, R.J., Fielding, D.J., Marland, G., Boden, T.A., & Kumar, N. Carbon dioxide emissions from fossil-fuel use, 1751-1950. *Tellus* 51B:759-65 (1999).
29. Kaufmann, R. K., Kauppi, H., Mann, M. L. & Stock, J. H. Reconciling anthropogenic climate change with observed temperature 1998–2008. *Proc. Natl. Acad. Sci.* **108**, 11790-11793 (2011).
30. Hansen, J.E. & Sato, M. Trends of measured climate forcing agents. *Proc. Natl. Acad. Sci.* **98**, 14778-14783 (2001).
31. Velders, G. J.M., et al. The importance of the Montreal Protocol in protecting climate. *Proc. Natl. Acad. Sci.* **104**, 4814-19 (2007).

32. Kai, F. U., Tyler, S. C., Randerson, J. T. & Blake, D.R. Reduced methane growth rate explained by decreased Northern Hemisphere microbial sources. *Nature* **476**, 194–197 (2011).

33. Hansen, J., Sato, M., Kharecha, P. & von Schuckmann, K. Earth's energy imbalance and implications. *Atmos. Chem. Phys.* **11**, 27031-27105 (2011).

34. Hansen, J., Sato, M., Lacis, A., Ruedy, R., Tegen, I. & Matthews, E. Perspective: Climate forcings in the industrial era. *Proc. Natl. Acad. Sci.* **95**, 12753-12758 (1998).

### **Acknowledgements**

FE acknowledges financial support for this work from the Consejo Nacional de Ciencia y Tecnología (<http://www.conacyt.gob.mx>) under grant CONACYT-310026, as well as from PASPA DGAPA of the Universidad Nacional Autónoma de México.

### **Author contributions**

F.E. and P.P contributed equally to the conceptual design, the data analysis and the writing of this manuscript. B.M.L Contributed to the conceptual design and data analysis.

### **Additional information.**

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at [www.nature.com/reprints](http://www.nature.com/reprints).

Correspondence and requests for materials should be addressed to F.E.

([feporrua@atmosfera.unam.mx](mailto:feporrua@atmosfera.unam.mx)).

**Competing financial interests**

The authors declare no competing financial interests.

## Figure Legends

Fig 1. Filtered and unfiltered temperature series. Panel a shows in dashed lines the observed G, NH and SH (left) and the filtered G and NH and unfiltered SH (right) for the NASA dataset, while the continuous lines represent the fitted temperature series using RFGHG, TRF\* and TRF. Panel b shows in dashed lines the observed G, NH and SH (left) and the filtered G and NH and unfiltered SH (right) for the HadCRUT4 dataset, while the continuous lines represent the fitted temperature series using RFGHG, TRF\* and TRF.

Fig 2. Aggregated radiative forcing series. Time series plot of RFGHG, TRF\* and TRF and the fitted trend functions with two breaks: 1960 and 1994 for RFGHG, 1960 and 1992 for TRF\*, 1960 and 1991 for TRF.

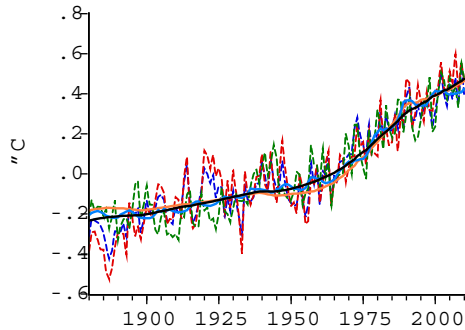
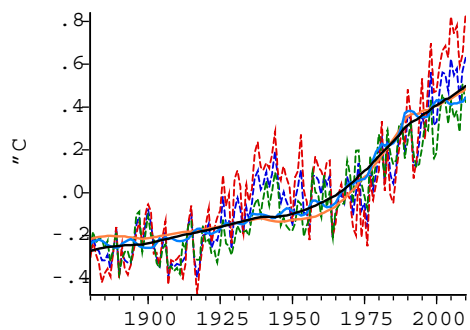
Fig 3. GDP and CO<sub>2</sub> emissions yearly changes. Panel a) shows the percent yearly changes in the GDP of Europe and of the USA, and of the CO<sub>2</sub> global emissions. Dark shaded areas mark the World Wars I and II, and the Great Depression. Light shaded areas are economic crises that took place within two years of these events. Panel b) depicts the yearly changes in the radiative forcing of CO<sub>2</sub> and RFGHG.

Fig 4. Trends of CFC<sub>11</sub>, CFC<sub>12</sub>, CH<sub>4</sub>, CO<sub>2</sub>, and AIE. Time series plot of some components of RFGHG and of AIE along with the fitted trend function with two breaks; the dashed lines indicate the dates of significant breaks: 1956 and 1993 for CFC<sub>11</sub>; 1957

and 1995 for CFC<sub>12</sub> (panel a); 1943 and 1992 for CH<sub>4</sub> (panel b); 1966 and 1996 for CO<sub>2</sub> (panel c); 1956 and 1988 for AIE (panel d).

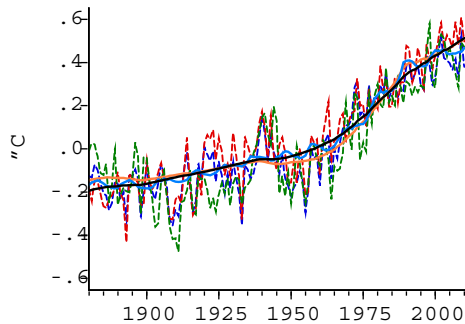
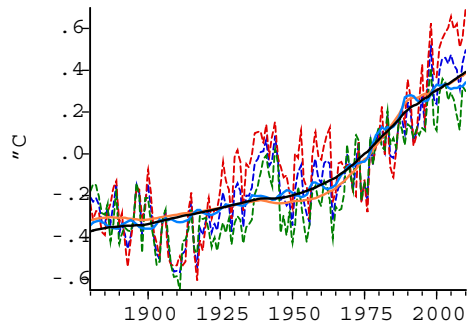
Fig 5. The effects of the slowdown in RFGHG and the increase in AIE over TRF. Time series plot of TRF, TRF without the break in RFGHG and TRF without the break in RFGHG and AIE.

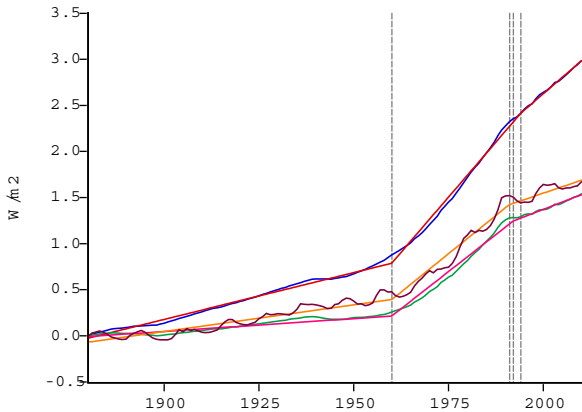
a



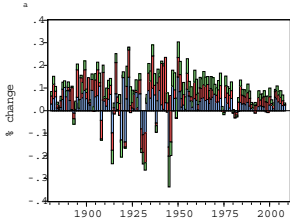
--- G  
--- NH  
--- SH  
--- TR F\* fitted  
--- TR F fitted  
--- RFGHG fitted

b

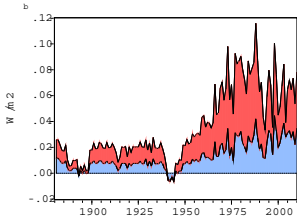




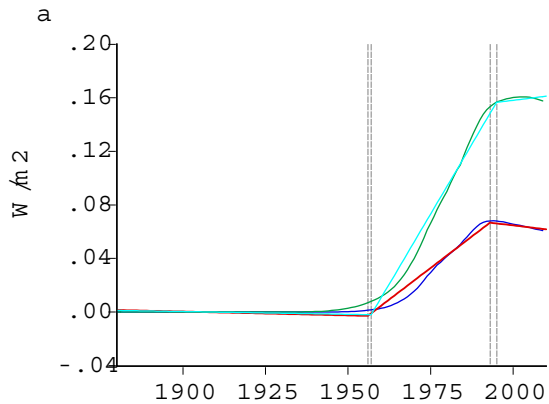
- RFGHG
- RFGHG fitted
- TRF\*
- TRF\* fitted
- TRF
- TRF fitted



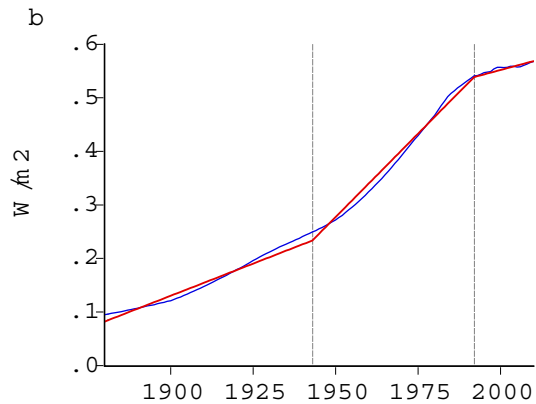
■ CO2 emissions  
■ GDP US  
■ GDP Europe



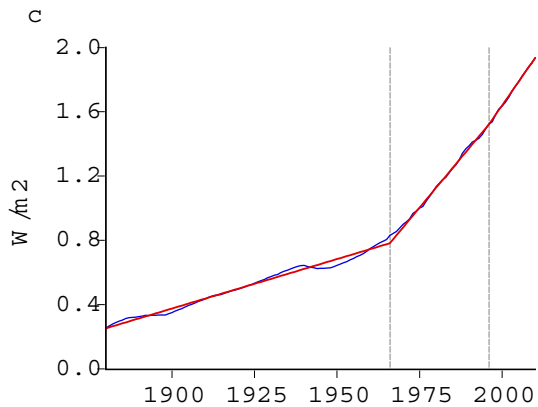
■ RFCO<sub>2</sub> ■ RFGHG



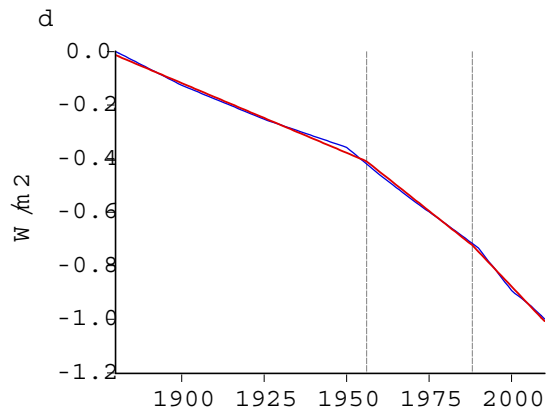
— CFC11  
 — CFC11 fitted  
 — CFC12  
 — CFC12 fitted



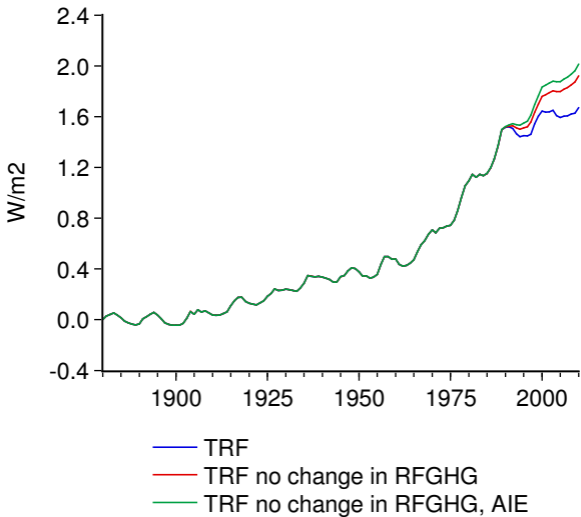
— CH4  
 — CH4 fitted



— CO2  
 — CO2 fitted



— AIE  
 — AIE fitted



## Statistically-derived contributions of diverse human influences to 20<sup>th</sup> century temperature changes

Francisco Estrada<sup>1,2</sup>, Pierre Perron<sup>3</sup> & Benjamín Martínez-López<sup>1</sup>

<sup>1</sup>*Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Ciudad Universitaria, Circuito Exterior, 04510 Mexico, DF, Mexico,* <sup>2</sup>*Institute for Environmental Studies, Vrije Universiteit, Amsterdam, Netherlands,* <sup>3</sup>*Department of Economics, Boston University, 270 Bay State Rd. Boston, MA, 02215, USA.*

### Supplementary Information

#### S1 Filtering the Atlantic Multidecadal Oscillation (AMO) from global (G) and Northern Hemisphere (NH) temperatures.

The AMO represents ocean-atmosphere processes naturally occurring in the North Atlantic with a large influence over the Northern Hemisphere and global climates. Recent publications have underlined the fact that these processes can mask the warming trend exaggerating it when the AMO is in its positive phase and masking it when in its negative<sup>1,2,3</sup>. Therefore, when conducting detection and attribution studies it is useful to control for the effects of AMO on G and NH in order to be able to extract a clearer externally forced signal. Here we present further empirical evidence on this issue. Note that although in the main text only the HadCRUT4 and NASA datasets are discussed, this Supplementary Information documents also includes the corresponding estimates from the HadCRUT3 dataset in order to provide a sensitivity analysis of our results.

As previously reported<sup>4,5,6</sup>, in general the results of applying standard unit root tests indicate that temperatures are unit root processes, but as discussed in S2.1 and S2.2 it has been shown that these tests are not adequate due to the existence of breaks in the trend function of temperatures (Supplementary Table 1). According to Supplementary Table 2, once a more adequate representation of the trend in temperatures is allowed, G, NH and the Southern Hemisphere (SH) temperatures are better described as trend stationary processes with a one-time break in the slope of their trend function. In the case of SH from NASA and of SH and G from HadCRUT4 there is evidence of another break (see S2.2, S2.3 and S2.4) at the beginning of the 20th century (Supplementary Table 2). For the HadCRUT3 dataset G, NH and SH show a single break and the estimate of the break date for SH is similar to that of the second break in NASA. Note that for SH from NASA and HadCRUT4 and G from HadCRUT4, the tests reported in Supplementary Tables 1-2 were applied to two subsamples in such a way that each of them contains only one structural break. It is worth noting that the unit root null is also rejected at the 5% level when using the whole sample, irrespective of the break date estimate used.

The estimated break dates occurring in the second part of the 20th century for G are 1976, 1971 and 1978, and 1982, 1984 and 1985 for NH, and 1976, 1955 and 1954 using the HadCRUT4, HadCRUT3 and NASA datasets, respectively. For all datasets, the estimated break dates for G and NH are very similar and not statistically different (Supplementary Table 3), being only a few years apart. In the case of SH, this is also true for the NASA and HadCRUT3 datasets but for SH from the HadCRUT4 the estimated break date occurs notably later. For SH, the estimate of the break date obtained from HadCRUT3 is not statistically different from that of NASA or HadCRUT4. However, the estimate of the break date of SH from HadCRUT4 is statistically different from that of NASA, although the intervals are only

separated by one year. As discussed below, these differences in the break date estimates of SH are produced by a low-frequency oscillation (similar to AMO but with a slightly different phase) that is only present in the HadCRUT4 dataset. The pre-break trend slopes of G, NH and SH show a slight warming of 0.28°C, 0.29°C and 0.31°C per century, respectively, for the HadCRUT4. The corresponding warming rates for the NASA dataset are 0.39°C, 0.42°C and 0.27°C. For the HadCRUT3 these figures are 0.22°C, 0.30°C and 0.11°C per century. The post-break warming rates are very similar between the three datasets, but as expected there are large differences between NH and SH due to the thermal inertia and feedback processes. The warming rates per century are 1.68°C, 2.76°C and 0.92°C per century for G, NH and SH, respectively, for the NASA dataset. The corresponding figures for the HadCRUT4 dataset are 1.58°C, 2.38°C and 1.24°C per century, and 1.51°C, 2.52°C and 1.11°C per century for the HadCRUT3 dataset.

Furthermore, the results of a nonlinear nonparametric cotrending test<sup>7</sup> indicate that G, NH, and SH share a common secular movement with the total radiative forcing (TRF) and the radiate forcing of greenhouse gases (RFGHG; Supplementary Table 4). As discussed in the main text and in section S3, this common trend is imparted by the radiative forcing, mainly determined by RFGHG. As such, to investigate the influence of AMO over G, NH and SH, we first detrended these temperature series by means of Ordinary Least Squares (OLS) regressions with TRF as the explanatory variable. Correlation analysis was applied to the residuals (G\_res, NH\_res and SH\_res) and to some known dominant modes of interannual climate variability. Both G\_res and NH\_res show an oscillating pattern very similar to that of AMO while this is not the case for SH\_res (Supplementary Figure 1). However, it is also noticeable that with the HadCRUT4, the influence of AMO on SH is found to be higher than with the HadCRUT3 and NASA datasets. The pair-wise correlation coefficients between G\_res, NH\_res and AMO are 0.75 and 0.79, respectively, for the HadCRUT4 dataset, 0.72 and 0.75 for the NASA dataset and 0.71 and 0.79 in the case of HadCRUT3. The correlation coefficients between AMO and SH\_res are notably lower for NASA and HadCRUT3 (0.41 and 0.38, respectively) but moderate for HadCRUT4 (0.54). Note that in all cases the interannual variations of temperature series around the underlying long-term trend are not as highly correlated with other prominent variability modes such as the Southern Oscillation Index (SOI), the North Atlantic Oscillation (NAO) and the Pacific Decadal Oscillation (PDO; Supplementary Table 5). These results strongly suggest the need to filter AMO from G and NH in order to reveal the underlying warming signal.

Since trends and breaks are low frequency features, it is important to purge the temperature series from natural low frequency components. This allows more precise estimates of the break dates. Other high frequency fluctuations in temperature series do not affect the precision of the estimates of the break dates and the magnitudes of the changes in slope. As shown in Supplementary Figure 1, detrended temperatures (G\_res and NH\_res) are correlated with the low frequency movements of the AMO. However, as shown in panels (d), (e) and (f), the filtered temperatures using equation (1) described below (GF\_res and NHF\_res) do not have long term movements and are affected only by high frequency components. The case of SH from the HadCRUT4 dataset is particular due to the existence of a low-frequency oscillation that is not present in the other SH series from other datasets and that is likely to produce the difference in the estimated break dates during the second part of the 20th century. The difference between SH from the HadCRUT4 and HadCRUT3 datasets (Supplementary Figure 2) reveals a low-frequency oscillation that has a change of phase from positive to negative starting in the mid-1950s that may affect the estimate of the break date moving it closer to its change to a positive phase occurring in 1976, exactly the estimated

break date of SH from the HadCRUT4. As shown in Supplementary Figure 2, the changes of phase during the second part of the 20th century of the difference between SH from the HadCRUT4 and HadCRUT3 and those of AMO match, although this is not true for the first half of the century. This explains the larger value of the correlation coefficient between SH from the HadCRUT4 and AMO when compared with those obtained using the HadCRUT3 and NASA datasets and also the differences in the estimated break dates. Importantly, if AMO is filtered from SH from the HadCRUT4 dataset using regression (1) below, the estimated break date is 1954 which is broadly similar to those of HadCRUT3 and NASA. This provides strong evidence about the effects of an AMO-like, low-frequency oscillation that is a particularity of the HadCRUT4 SH series affecting the break date estimates.

We use OLS regressions to filter AMO from G and NH as follows:

$$T_t = c + b \text{AMO}_t + T_t^* \quad (1)$$

where  $T_t$  is the unfiltered temperature time series,  $c$  and  $b$  are unknown but fixed parameters and  $T_t^*$  is the filtered temperature time series (the residuals of the regression). The objective of this procedure is not comparing results from different filtering schemes, but to have the low frequency component purged as much as possible from low frequency natural effects. Notice that potential measurement errors in AMO have no impact on the results since they cannot affect the trend and, hence, have no impact on the tests (except possibly to affect power but since the tests for breaks and unit root reject this is not an issue). The measurement errors may have an impact on the temperature high frequency variability but here no attempt to model the movements around the trend is carried out.

Supplementary Figures 3, 4 and 5 show the unfiltered and filtered G and NH and the unfiltered SH temperature series for the HadCRUT4, the NASA and the HadCRUT3 datasets. Once the effects of AMO are filtered from the temperature series, a nonlinear warming trend that is similar for all G, NH, SH series is unveiled in all cases. A salient characteristic of these nonlinear trends is an abrupt change in the rate of growth occurring in the mid 20th century. The results of the statistical tests supporting the existence of these breaks are presented in S2.2.

Note that before conducting the filtering procedure, three standard unit root tests<sup>8,9,10</sup> were performed on AMO in order to insure that in doing so no trending behavior could be spuriously induced to or taken away from temperature series. As shown in Supplementary Table 6 all standard tests provide evidence for concluding that AMO can be considered stationary around a constant mean.

Furthermore, the Bierens nonparametric nonlinear co-trending test<sup>7</sup> was applied to both unfiltered and filtered G and NH series in order to test whether filtering AMO from G and NH using regression (1) affects the underlying trend of these temperature series or only removes part of their low-frequency variability as intended (see also S3). The results shown in Supplementary Table 7 indicate that the unfiltered and filtered versions of G and NH share the same nonlinear trend and thus the filtering procedure does not alter their secular trend.

Note that all results remain the same whether the filtered or unfiltered data are used (Supplementary Tables 1 to 4, 7, 9 and 11) with the exception of the estimates of the break dates and the magnitudes of the changes in the slopes. The break date estimates for the filtered G and NH temperature series are close to those for the unfiltered SH (NASA and

HadCRUT3) and the forcings, in agreement with what one can expect from the physics of climate change (see S7). The break dates for the unfiltered G and NH temperatures series occur much later due to the fact that the AMO entered its negative phase around the beginnings of the 1950s, masking the warming trend until the mid 1970s when the positive phase of the AMO started.

## **S2. Statistical analysis of temperatures and radiative forcing series.**

### **S2.1. Standard unit root tests applied to temperature and radiative forcing series.**

A first step to investigate the univariate time-series properties of G, NH (filtered) and SH, as well as those of RFGHG, TRF\* and TRF, consists in applying standard unit root tests that are commonly used<sup>8,9,10</sup>. The tests results clearly show a failure to reject the null hypothesis of a unit root in all the temperature and radiative forcing series. This is similar to results previously reported for observed global and hemispheric temperature series and for radiative forcing variables<sup>4,5</sup>. However, it has been shown that if structural breaks are present in the trend function of the time series under investigation and are not taken into account, unit root tests can be severely biased towards a non-rejection of the unit root hypothesis<sup>11</sup>. Previous analyses have offered evidence about the existence of structural breaks in both observed and simulated temperatures, and more recently in the radiative forcing series<sup>6,12,13</sup>. Such studies have shown that the results obtained by applying standard unit root tests are reversed when the occurrence of a one time-break in the trend function is allowed. Visual inspection of both the filtered G, NH and the unfiltered SH suggests that the results of the standard unit root tests can be influenced by an incorrect specification of the trend function. Thus, formal testing for the presence of potential breaks must be conducted in order to investigate the time-series properties of these variables. As shown below, the existence of a break in the slope of the trend function of all temperature and radiative forcing variables is confirmed by recent structural change tests that are valid for both integrated and stationary noise components.

In addition, the results of applying the same standard unit root tests to the radiative forcing of stratospheric aerosols (SAER) show that SAER is a stationary process around zero, without any noticeable structural breaks in its intercept or slope coefficients. As such, although this variable produces important short-lived effect on climate, given its time-series properties it cannot impart any long-term trending behavior to temperature series. For this reason it is not included in TRF and TRF\*. Furthermore, consider the following statistically adequate regression model for G from NASA<sup>13</sup>:

$$T_t = -0.15 + 0.28TRF_t + 0.05SAER_t + 0.26T_{t-1} - 0.03SOI_t + 0.33AMO_t + \varepsilon_t$$

where all estimated coefficients are statistically significant at the 5% level and the coefficient of determination of the regression ( $R^2$ ) is 0.93. Estimates of how long it takes for a shock to dissipate can be computed using the coefficients of this regression. The half-life of a shock to SAER on temperatures is about 6 months, while it would take around 3.5 years for 99% of the effects to dissipate. Broadly similar results were obtained for HadCRUT4 and HadCRUT3. This provides further evidence supporting that SAER cannot impart a long-term trend to temperature series.

### **S2.2. Structural change test when the order of integration of the noise component is unknown.**

The Perron-Yabu<sup>14</sup> procedure to test for the presence of a structural change in the trend was applied to both temperature and radiative forcing series. This test has the advantage over other available tests for change in trend in that it is valid whether the noise component is integrated of order one or stationary, thus circumventing the circular problem of pre-testing for unit roots. The test statistics shown in Supplementary Table 8 are significant at the 1% level for all temperature and radiative forcing series, providing substantial evidence for the existence of a break in the slope of their trend function. Thus, the results of standard unit root tests can be misleading and unit root tests that allow for a one-time break in the trend function are required to investigate the type of data generating process that best describes temperature and radiative forcing variables.

### **S2.3. Unit root tests that allow for a one-time break in the trend function.**

It has been shown that the estimate of the sum of the autoregressive coefficients is highly biased towards unity if there is a shift in the trend function<sup>11</sup>. In this case, the unit root null is hardly rejected even if the series is composed of *i.i.d.* disturbances around a trend. Furthermore, if the break occurs in the slope of the trend function, unit root tests are not consistent, i.e., the null hypothesis of a unit root cannot be rejected even asymptotically<sup>11</sup>.

To analyze the time-series properties of temperature and radiative forcing variables we applied the Perron<sup>15</sup> and Kim-Perron<sup>16</sup> tests using the "changing growth" model specification, which allows for a one-time change in only the slope of the trend function<sup>11</sup>. Both unit root tests allow for a structural change occurring at an unknown date, which is estimated by minimizing the sum of squared residuals from regression (1) in the main text.

Note that, although visual inspection of some of the temperature series could suggest the existence of a level shift in addition to the change in the slope, the changing growth model specification is preferable for the following reasons: 1) there is no concurrent level shift at the time of the change in slope for the forcings, hence one should not expect a permanent level shift in temperatures. Given the large variability in temperatures, some random variations can appear as a level shift in trend; 2) more importantly, as shown in the econometrics literature<sup>22</sup>, the rate of convergence of the estimate of the break date is faster when no concurrent level shift is allowed. In addition, if there is a concurrent level shift of small to moderate magnitude, it is better to ignore it and use a trend function with only a slope shift<sup>22</sup>. This allows a more precise estimate of the break date.

The Kim-Perron<sup>16</sup> test requires pretesting for the existence of the break using some structural change test such as the Perron-Yabu<sup>14</sup> test applied in S2.2. Given that this pre-test rejects the null of no structural change, the limit distribution of the unit root test is then the same as if the break date was known and the test has much greater power<sup>11,17</sup>.

The results in Supplementary Table 9 indicate that the unit root null hypothesis can be strongly rejected for all temperature and radiative forcing series: for the temperature series the test statistics are significant at the 1% level and for all of the radiative forcing series they are at 5% level. What these results indicate is that the non-rejection of the unit root hypothesis using standard unit root tests is caused by an incorrect specification of the trend function as also demonstrated in recent publications<sup>6,12,13,18</sup>. A direct implication of this finding is that cointegration techniques are not adequate to investigate the attribution of climate change (i.e., the long-run relationships between temperatures and radiative forcing series) because this technique relies on the assumption of stochastic trends. As shown in the econometric literature, under similar circumstances spurious cointegration is likely to occur

and the inferences drawn are not reliable<sup>19,20,21</sup>. As a consequence, a different approach to test for common long-run relationships is required.

#### **S2.4. Constructing confidence intervals for the break dates.**

The break dates presented in the previous sections are only point estimates making it difficult to assess how different or similar they really are. To investigate this issue we applied the Perron-Zhu<sup>22</sup> methodology to construct 95% confidence intervals for these parameters. As shown in Supplementary Table 10, the break date estimates for all temperature and radiative forcing variables are not statistically different from each other (except for SH from HadCRUT4 as discussed previously). Hence, the Perron-Zhu<sup>22</sup> procedure indicates that the abrupt change in the rate of growth of all temperature and radiative forcing series documented in the previous sections occurred at a date that can be considered common to all series. Consequently, the results in Supplementary Tables 9 and 10 suggest the existence of a secular co-movement between global and hemispheric temperatures and radiative forcing variables, which is consistent with climate physics<sup>23,24,25</sup>. The existence of this common long-run path is investigated in the following sections by means of statistical tests that are adequate for the time-series properties that have been shown to best describe these variables.

#### **S3. Testing for nonlinear co-trending in temperature and radiative forcing variables.**

Cointegration techniques<sup>26</sup> are commonly applied to study the attribution of observed climate change, although the basic assumption needed for this procedure (i.e., variables that are integrated processes) was never soundly tested for temperature variables until recently. Detailed analyses of the time-series properties of these variables have offered evidence against unit root processes providing an adequate representation for both observed and simulated global and hemispheric temperatures<sup>6,12,13,18</sup>. As shown in this paper the assumption of a unit root is neither adequate for radiative forcing series.

However, it is important to take into account that unit root processes are not the only type of nonstationary processes that can show a common secular movement and that cointegration analysis is only one possibility for relating the trends of nonstationary variables. Relationships between nonstationary variables can be established when linear combinations of different time series cancel out some "common features" such as trends and breaks<sup>27</sup>.

Supplementary Figures 6a to 6c show the fitted temperature series obtained from OLS regressions with RFGHG, TRF\* and TRF, respectively, as the explanatory variable with the filtered G and NH, and the unfiltered SH as the dependent variables, along with the corresponding residuals. The concordance between the forcing and temperature trends is apparent; in all cases the fitted values appear the product of a low-pass filter on temperatures. Visual inspection of the residuals in Supplementary Figure 6a to 6c strongly suggests that in all cases all secular movements have cancelled, leaving only stationary variations around zero.

In this section we apply the nonparametric nonlinear co-trending analysis proposed by Bierens<sup>7</sup>. Nonlinear co-trending is a special case of common features in which one or more linear combinations of possibly nonlinear trend stationary time series are stationary, indicating that the series share common nonlinear deterministic time trends.

Supplementary Table 11 shows the results of applying this test to three sets of time series composed of a) the filtered G, NH and unfiltered SH temperature series from the

HadCRUT4 dataset, RFGHG and TRF; b) the filtered G, NH and unfiltered SH temperature series from the HadCRUT3 dataset, RFGHG and TRF; c) the filtered G, NH and unfiltered SH temperature series from the NASA dataset, RFGHG and TRF. In all cases the results point to the existence of four co-trending vectors ( $r = 4$ ) involving G, NH, SH, RFGHG and TRF, hence to the existence of one common nonlinear deterministic trend. Therefore, it is possible to express the nonlinear trends in four variables as constants times the nonlinear trend in the remaining variable. Choosing RFGHG as the numeraire the common nonlinear trends can be expressed as, say  $a+b*\text{RFGHG}$ , where  $b$  is given by:

	HadCRUT4 dataset	HadCRUT3 dataset	NASA dataset
G	0.1605	0.2569	0.1816
NH	0.1764	0.2091	0.1659
SH	0.1448	0.3048	0.1968
TRF	0.6366	0.6364	0.6361

The coefficients for G, NH, SH are similar within datasets but are systematically higher for the HadCRUT3 dataset, particularly in the case of SH. Notably, the estimates from the HadCRUT4 are more similar to those obtained using NASA than to the previous version of the same dataset. These differences are likely due to the large changes made in the HadCRUT4: major update of the sea-surface temperature component, the inclusion of newly digitised data measured over both sea and land and re-homogenized station data. The coefficient relating TRF and RFGHG points to the fact that the effect of all other forcing factors has been to reduce TRF without altering the general trend imparted by RFGHG, which accounts for most of the human contribution to global warming. The results, interpreted according to climate physics, suggest a dominant anthropogenic influence on observed warming: the simplest radiative forcing series considered in each group of variables is RFGHG which contains the nonlinear trend that is present in all the other series, suggesting that this variable is the dominant driver imparting the common nonlinear trend to TRF and in turn to G, NH, and SH.

#### **S4. The effect of World War I (WWI), the Great Crash and World War II (WWII) on GDP, CO<sub>2</sub> emissions, radiative forcing and global and hemispheric temperatures.**

Note that although the description of temperature and forcing variables as having piecewise linear trends is convenient to investigate their time-series properties, it is a large simplification of the much more complex secular movement of these variables. What the analyses so far revealed is that the overwhelming feature of the data is a large change in the rate of increase near 1960, that variations are stationary around these broken trends and that there is a common trend with a break in radiative forcing and temperature variables; the break need not be abrupt, a smooth transition is possible. This does not preclude other nonlinearities that could not be detected by the unit root tests on the residuals, such as a transition period or other small breaks. In fact, this is strongly confirmed by the results of the nonlinear nonparametric co-trending test<sup>7</sup> which show that both temperatures and radiative forcing variables share a common nonlinear secular movement which is driven by RFGHG. One feature of observed global and hemispheric temperatures that is commonly referred to in the literature is the cooling period of the 1940-1970's<sup>28,29,30,31,32</sup>. This has been explained as a combination of natural variability and the negative radiative forcing from the reflective tropospheric aerosols (RAER) produced during WWII. While both of these factors may have been contributors to the cooling period, the CO<sub>2</sub> radiative forcing (RFCO2) is another cause that has rarely been mentioned in the discussion so far. RFGHG from 1938 to the early 1950's

experienced a considerable slowdown in its rate of growth, remaining almost completely flat from 1938 to 1947 and then experienced a rapid and sustained increase which is marked by the occurrence of the break in the slope of its trend function around 1960 (Supplementary Figure 7a, Supplementary Tables 8 and 10). That is, the lack of increase in RFGHG is another factor that contributed to this cooling period, and it was mainly produced by a decrease in RFCO<sub>2</sub> for almost a decade (1940s, Supplementary Figure 7b). Otherwise, the cooling effect of RAER would have been mostly compensated by the increase in RFGHG: the yearly rate of growth of RAER during this cooling period was -0.013 W/m<sup>2</sup> while that of RFGHG before the slowdown in RFCO<sub>2</sub> was 0.010 W/m<sup>2</sup>.

The percent reduction in CO<sub>2</sub> emissions during the period 1914-1946 has no parallel with any other period since 1751. These reductions were driven by three landmark socioeconomic events that occurred in a short period, making it hardly comparable to any other period in recent history: the two World Wars and the largest and longest economic crisis until now. To account for the global economic activity during the 20th century, we use the GDP of the United States of America and of the 12 most important economies in Europe and as expected a strong relationship was found between the growth rates of GDP and CO<sub>2</sub> emissions (highly significant correlation coefficients of 0.51 and 0.68 between CO<sub>2</sub>, USA and Europe GDP, respectively).

As mentioned above, these socioeconomic disruptions caused large reductions in CO<sub>2</sub> growth rates that are uncommon in the context of CO<sub>2</sub> emissions since 1751. In all the occasions since 1751 for which there were negative yearly percent changes in this variable, about half of a total of 27 decreases cluster in the 1914-1946 period and it was also during this period that the largest decreases since the beginning of the Industrial Revolution occurred (up to 16%). Although other factors in the global carbon cycle may have contributed, CO<sub>2</sub> emissions suggest that these reductions led first to a decrease in the rate of growth in RFCO<sub>2</sub> and consequently in RFGHG for almost four years starting in 1937, to negative growth rates for another four years (1941-1944). It was not until 1949 that RFCO<sub>2</sub> and RFGHG regained their previous growth rate levels (Figure 3b in the main text).

For all (HadCRUT4, NASA and HadCRUT3), visual inspection of G, NH and SH suggests the existence of a "pause" in the warming around the mid-20th century. Fitting a linear trend by means of OLS to the HadCRUT4 data over this period (1940-1970) results in negative but not statistically significant slope parameters of -0.024°C, -0.030°C and -0.017°C per decade for G, NH and SH, respectively. When the influence of AMO has been filtered from G and NH, and therefore the effects of the negative phase of this oscillation have been taken out, the pause in warming is still suggested by visual inspection, but with a shorter duration approximately extending from 1938 to 1955. The values of the slope parameters during this period are about -0.07°C, -0.06°C and -0.07°C per decade for G, NH and SH from HadCRUT4. Note, however, that there is no statistically significant evidence for changes in the slope parameters of the temperature series. This pause in the rate of growth is also present in the nonlinear trend of RFCO<sub>2</sub> and RFGHG.

#### **S5. The anthropogenic contribution to the current slowdown of global warming: an analysis of the components of RFGHG.**

According to different sources there is growing evidence to support the fact that during the last two decades the rate of warming has decreased<sup>33,34,35</sup>. As a result, a discussion about the possible origin of a slowdown in global warming since the 1990's has taken place<sup>34,35</sup>.

Internal natural variability, particularly El Niño/Southern Oscillation (ENSO) and variations in solar forcing as well as the effect of RAER have been proposed as the main possible factors for having cooler than previously expected annual global temperature records. It has also been suggested that this slowdown in warming does not represent a change in the underlying trend but that it is mainly driven by the sum of temporary natural variations. As such, the perception of the slowdown in the warming could largely disappear as more years of data are added<sup>33</sup>.

In this section we present evidence supporting the fact that in addition to natural variability, solar forcing and RAER, the current slowdown is in great part the outcome of large changes in anthropogenic greenhouse forcing. These changes have been large enough to produce a highly statistically significant structural change in the slope of the trend function of RFGHG, TRF\* and TRF and therefore should not be expected to revert in the short-term unless large and sustained additional increases in external forcing occur.

In particular, we show that the effects of the Montreal Protocol and of changes in agricultural practices in Asia have been large enough to change the long-run path of radiative forcing which, as documented in the previous sections, drives global and hemispheric temperatures. Furthermore, these breaks in the rate of growth of radiative forcing provide evidence on the actual efficacy of reducing GHG gases other than CO<sub>2</sub> for slowing global warming in the short-term.

To obtain statistical evidence about the causes of the slowdown in the increase in temperatures, we search for additional breaks in the post 1960 period using the Kejriwal-Perron<sup>36</sup> sequential testing procedure based on the Perron-Yabu<sup>14</sup> test. As shown in Supplementary Table 12, the slowdown in RFGHG is confirmed by a highly significant change in its trend occurring in 1994 producing a 25.61% reduction in its slope. This break is even more pronounced when the direct and indirect effects of tropospheric aerosols are considered, as in TRF\* and TRF. For these two series, the breaks are also highly significant, and occur in 1992 and 1991, respectively. Their corresponding rates of growth were reduced by more than 50% in both cases. Note that the estimated break dates are not statistically different from each other. In addition, when considering the 1960-2010 sample, there is evidence of a statistically significant break in the slope of the filtered G in 1992 (Sup-Wald test for NASA and Chow tests with this break date for NASA and HadCRUT3 are significant at the 5% level) implying a decrease in the rate of growth of temperatures. This evidence is strong in view of the fact that the sample is quite short. It should be noted, however, that this result is not robust to using a longer sample.

As documented by the results shown in Supplementary Tables 8 to 11 the change in the growth rate of radiative forcing is a common feature of the forcing and temperatures trends. Furthermore, when detrending temperature series using RFGHG, TRF\* or TRF all of the residuals are stationary and no sign of this structural change nor that occurring in the period 1955-65 reported above could be found when applying a unit root test or the Perron-Yabu<sup>14</sup> procedure from the beginning of the 20th century.

The decrease in the growth rate of RFGHG can be traced back to the reduction of CFCs emissions and to the "pause" in the growth rate of atmospheric methane<sup>37,38,39,25</sup>. On the one hand, the decrease in CFCs emissions was produced by the entrance into force of the Montreal Protocol in 1989 for controlling the substances that deplete the stratospheric ozone layer. On the other hand, the decrease of atmospheric methane is not completely understood

but appears to be related to the decrease in microbial sources caused by the application of chemical fertilizers and to a more efficient water use for producing rice in Asia<sup>39</sup>.

To obtain a better assessment of the causes of the slowdown in temperatures, we applied the sequential break point detection procedure of Kejriwal and Perron<sup>36</sup> to all the components of RFGHG using two samples: pre 1990 and post first break date. We use this subsample approach because, as documented in the structural change literature, a test for one break will have difficulty detecting the existence of a break when in fact two breaks are present and are characterized by an increase followed by a decrease. The subsample approach avoids this problem and leads to tests with higher power. This led to the conclusion that two breaks are present for all components, except for N<sub>2</sub>O for which a single break is present. However, only four of the RFGHG components showed statistically significant evidence for a break in the post 1990 period. Given the potential contribution of RAER, AIE and BC to the slowdown in the warming, the existence of breaks in the post 1960 period in these series was also investigated. AIE is the only variable for which a significant second break could be detected. The results are shown in the Supplementary Table 12.

The largest reduction occurred for CFC<sub>11</sub>, which has a break in 1993 with the slope decreasing almost 116%, a reversal to a sustained decrease. The second in importance occurred for CFC<sub>12</sub> with a break in 1995 and a decrease in slope of 92.85%, i.e., inducing a nearly zero rate of increase. These results provide clear evidence that the Montreal protocol was indeed successful in achieving global reductions in CFC emissions. Although it was not its objective, the reductions were large enough to have an impact on global RFGHG, TRF\* and TRF, which slowed down the increase in the warming<sup>38</sup>. The third most important factor is CH<sub>4</sub>, which has a break in 1992 with the slope decreasing 73.35%, which is about 30% lower than its growth rate at the beginning of the century. The only other series included in RFGHG with a break in the 90s is CO<sub>2</sub> but it actually exhibits an increase in its slope of 19.78% in 1996, indicating that the efforts undertaken until now have not been sufficient to decrease the radiative forcing associated with CO<sub>2</sub>. Hence, the evidence points to the fact that the decrease in CFC<sub>11</sub>, the flattening of CFC<sub>12</sub> and of CH<sub>4</sub>, all of them having large global warming potentials, are major contributors for the decrease in temperatures during the last two decades, despite the more rapid increase in CO<sub>2</sub>.

Estimates show that if the second break in the RFGHG would have not taken place, the radiative forcing in the year 2010 would have been about 0.25 W/m<sup>2</sup> larger, an amount equivalent to the full amplitude of the solar forcing. The individual contributions are about 0.13 W/m<sup>2</sup>, 0.05 W/m<sup>2</sup> and 0.08 W/m<sup>2</sup> for CH<sub>4</sub>, CFC<sub>11</sub> and CFC<sub>12</sub>, respectively. Using the slope estimate  $\beta$  from the regression  $G_t = c + \beta TRF_t + \varepsilon_t$ , this difference in RFGHG would have made global temperatures about 0.1°C warmer in 2010. Global temperature anomaly in 2010 would have reached 0.73°C, 0.65°C and 0.59°C in the case of the NASA, HadCRUT4 and HadCRUT3 series, respectively. Accordingly, 2010 would have been undisputedly the hottest year in record, given the realizations of the other external forcing and natural variability factors that occurred, estimated using the deviations from the fitted trend.

The recent slowdown in warming can be seen as the product of an ordered sequence of structural changes induced from human systems to the emission of greenhouse gases and tropospheric aerosols, to the atmospheric concentrations/radiative forcing and finally to the global and hemispheric temperatures.

Paradoxically, the decrease in the warming rate of the last two decades, which is presented by global warming skeptics as a proof that human cannot affect the climate system, is shown to have a profound human origin.

### **S6. Zero dimensional energy balance models, cotrending and trend-stationarity.**

In this section, the co-trending time-series models used for the attribution of climate change are interpreted in terms of zero dimensional energy balance models. Previous publications have laid out the structural foundations of time-series models applied to the analysis of global mean surface temperature series<sup>40,41</sup>. The time-series models presented in this paper have the general form:

$$T_t = \alpha + \gamma F_t + \varepsilon_t \quad (2)$$

where  $T_t$  is temperature,  $F_t$  is some measure of radiative forcing and  $\varepsilon_t$  is a climate noise that encompasses both short-term variability and longer-term variability modes. The structural model behind this statistical model can be described by means of a simple two-compartments climate model<sup>40,42,43</sup>. The upper compartment is composed mainly of the atmosphere and the upper ocean and has a small heat capacity and short time constant for reaching its steady state. This upper compartment is thermally coupled to the lower compartment, composed of the deep ocean, which has a large heat capacity and long time constant for reaching the steady state. When a positive and sustained external forcing is applied to the system, the upper compartment temperature increases inducing changes in the absorbed and/or emitted radiation at the top of the atmosphere and a heat flow to the lower compartment, which has a much larger heat capacity and requires much longer time to respond to any forcing. This two-compartments model can be described by the following equations representing its upper and lower components, respectively<sup>40</sup>:

$$C_U \frac{dT_U}{dt} = F - \lambda \Delta T_U - \beta (\Delta T_U - \Delta T_L)$$

$$C_L \frac{dT_L}{dt} = \beta (\Delta T_U - \Delta T_L)$$

where  $\lambda$  and  $\beta$  are the climate response and the heat exchange coefficients, respectively. The time constants that characterize the response of these two compartments are very different, having been estimated to be about from 4 to 9 years for the upper compartment and about 400-580 years for the lower compartment<sup>40,43</sup>. The analyses presented in this paper relate to the response of the upper compartment of the climate system to increases in radiative forcing. The transient climate sensitivity characterized by the short time constant of the upper compartment is given by  $S_{tr} = (\kappa + \lambda)^{-1}$  where  $\kappa$  is the heat uptake coefficient of the climate system<sup>40</sup>. The transient climate sensitivity relates the time-dependent increase in surface temperature to the time-dependent forcing such that  $\Delta T(t) = S_{tr} F(t)$  and it is equal to the slope parameter  $\gamma$  in the time-series model described by equation (2) above. The response of the climate system to the forcing over the observed period is determined by the time constant of the upper compartment and the transient climate sensitivity, providing a physical explanation of why global and hemispheric surface temperatures follow the same nonlinear trend depicted by the radiative forcing (co-trending) and why observed temperatures rapidly

adjust to changes in the trend of the radiative forcing. These features are particularly clear after part of the low-frequency oscillations produced by the AMO are filtered out of both global and northern hemisphere surface temperatures. The overlapping confidence intervals in the break dates in the slope functions of radiative forcing and temperature series found in this paper are also consistent with the short time constant dominating this relationship, giving physical support to the idea of co-breaking in surface temperature and radiative forcing series.

As depicted in Supplementary Table 13, the estimates of  $S_r$  from different global surface temperature series, with or without filtering the AMO, and TRF provide broadly similar estimates that are also consistent with the ones that have been reported in the literature based on observational datasets<sup>40</sup>. This result also supports the working hypothesis that the effects of the AMO can be removed by a simple regression without affecting the long term-trend of temperatures and that the AMO can be described as part of the natural variability component  $\varepsilon_t$ .

The breaking trend stationary process used to represent the aggregated radiative forcing can also be explained in terms of a structural model. Any time series description of radiative forcings and temperatures is obviously a simplified version of a more complex dynamical system. It is nevertheless useful to characterize the salient features of the data. As argued above by means of a simple two-compartments climate model<sup>40</sup>, the time-dependent change in global surface temperature is proportional to the time-dependent radiative forcing  $\Delta T(t) = S_r F(t)$  but these changes are relatively insensitive to the nature of the forcing. As such, for the attribution problem addressed in this paper the crucial aspect is the modeling of the common trend.

The nature of the trend in the aggregated measure of radiative forcing can be thought of using the following decomposition of the time series properties of the  $i$ th forcing,  $y_{i,t}$ :

$$y_{i,t} = f_{i,t} + u_{i,t}$$

where  $f_{i,t}$  is the trend and  $u_{i,t}$  are deviations from the trend. The trend is of general form (including the nonlinear case) and intends to capture the sustained increase in the  $y_{i,t}$  forcing.  $u_{i,t}$  represents the noise component. Shocks to short-lived radiative forcings (ranging from aerosols to methane, for example) will dissipate shortly, while in the case of long-lived forcing (i.e., CO<sub>2</sub>, N<sub>2</sub>O, CFCs) the effect of shocks will be long-lasting. The overall trend of the aggregate forcing is the sum of the individual trends of each forcing in the set. The same applies to the noise components. Since CO<sub>2</sub> is by far the most important forcing, the overall trend will be substantially influenced by the trend in CO<sub>2</sub>.

The matter of interest related to whether the overall series are trend stationary or integrated (i.e., having an autoregressive unit root) refers to the nature of the deviations  $u_{i,t}$  for the components having a long atmospheric residence time, namely CO<sub>2</sub>, N<sub>2</sub>O, and CFCs, which have a so-called life-time in excess of 45 years, defined as the period it takes for a perturbation to be reduced to 37% of its initial amount. Given the time span of the historical records under study, we can for all practical purposes view the effect of such deviations as

having a permanent effect. This would lead to each of the long-lived forcings as having a unit root representation and hence the aggregated forcing would be integrated of order one.

The central issue then is the nature of the deviations for the long-lived components. To analyze this issue first note that all components of RFGHG except for CH<sub>4</sub> are long-lived. The radiative forcing of the long-lived radiative factors (RFL composed of RFGHG minus CH<sub>4</sub>) is shown in panel a) of Supplementary Figure 8, along with a fitted trend with a one-time break in 1960. The distinctive feature is that the series is very smooth and follows very closely the trend function. That is, the deviations are for all practical purposes negligible. Note that the effect of the sustained reductions in CFC<sub>11</sub> and CFC<sub>12</sub> in the 90s are present but minor given that CO<sub>2</sub> exhibited a sustained increase at roughly the same time (recall that the reduction in CH<sub>4</sub> is not included in this series of aggregated long-lived forcing). Besides other minor non-linearities, this is true except for one episode, namely the slowdown in RFGHG during the period 1940-driven by unprecedented reductions in CO<sub>2</sub> emissions during the period 1914-1946 linked to three landmark socioeconomic events that occurred in a brief period: the two World Wars and the Great Crash.

This slowdown can be viewed as a shock, though it occurred over several years and, hence, imparted an inflection point in the trend of WM-GHG instead of a sudden level shift. It nevertheless had a permanent effect. Supplementary Figure 8 panel b) presents a counterfactual experiment in which the slowdown did not occur. It assumes that the same trend continued until 1950 and that afterwards a trend with slope given by the post-1950 data was in effect. This shows that, indeed, if the slowdown during the period 1940-1950 had not occurred, the level of concentrations would have been permanently higher in the post-1950 period. So the slowdown can indeed be viewed as a shock (a deviation from the trend) that had a permanent effect in lowering the level of the aggregate radiative forcing.

The crucial matter of interest is that the shock, or deviation from trend, of the 1940-1950 period was the only non-negligible one that occurred in the sample, no important deviations from trend occurred in RFCO<sub>2</sub>, or RFGHG in general, at any other time. In terms of the statistical model, the shocks to concentrations  $u_{i,t}$  for those long-lived components have been for all practical purposes zero throughout the sample, except for the slowdown in 1940-1950 and the break in CFCs in the 1990s. Hence, from a statistical point of view, it is more adequate to view the process describing long-lived forcing as being a pure trend with an inflection point in the mid-20th and a marked change in slope in 1960. When adding all other forcing factors with a short lifetime, for which deviations from trends having a transitory effect have occurred, this leads to a segmented trend model with stationary deviations (no unit root) for RFGHG, TRF\* and TRF as depicted in Fig. 2 in the main text.

### **S7. Sensitivity to different radiative forcing for the direct effect of aerosols.**

Atmospheric aerosols are the largest contributors to the uncertainty in radiative forcing during the industrial period and the level of scientific understanding associated with this forcing is medium to low<sup>45,25</sup>. Aerosols have a dominant role in the climate change observed during the industrial period, and the results of any study on attribution, climate sensitivity and model performance evaluation may be sensitive to the exact specification of this factor<sup>40,45,46</sup>. The uncertainty in the radiative forcing from atmospheric aerosols has been estimated to be roughly of the same magnitude as the radiative forcing from greenhouse gases<sup>45,46</sup>. This section presents a sensitivity analysis of the results presented in this paper to an alternative series of the radiative forcing from atmospheric aerosols (direct effect).

For the sensitivity analysis, the total direct aerosol forcing (RAER\_SENS) available from the Representative Concentration Pathways (RCP) emission scenario database<sup>47</sup> is used to replace RAER in TRF. This new aggregate forcing series will be referred henceforth as TRF\_SENS. As in the case of TRF, the standard unit root/stationarity tests support the existence of a unit root in TRF\_SENS. Nevertheless, as revealed by the Perron-Yabu<sup>14</sup> test (significant at the 1% level, test statistic value of 6.62), the trend function of this series is also characterized by a large break in its slope occurring in 1962. As in the case of TRF and G, NH and SH temperatures, once the specification of the trend function is adequate (i.e., the break is incorporated) the unit root null is rejected at the 5% level (Supplementary Table 14).

The Bierens nonparametric nonlinear co-trending test<sup>7</sup> indicates the existence of four co-trending vectors and thus a common secular movement amongst G, NH, SH, TRF\_SENS and RFGHG for all cases with the HadCRUT4, NASA and HadCRUT3 datasets (Supplementary Table 15). This result also holds when AMO is filtered from G and NH (Supplementary Table 16).

Atmospheric aerosols have been suggested as one of the most important factors behind the pause in the warming during the 1940-1970 period and the recent slowdown<sup>35,25</sup>. Accordingly, we use RAER\_SENS to assess the adequacy of our results with respect to the influence of RFGHG during these two periods. As shown in Supplementary Table 17, there is statistical evidence of two breaks in the slope of the trend of RAER\_SENS during the 20th century, occurring in 1944 and 1987, close to both periods of interest (Supplementary Figure 9). In the first case, the slope of the radiative forcing of RAER\_SENS changes from being practically zero (-0.072 W/m<sup>2</sup> per century) to -0.061 W/m<sup>2</sup> per decade. The average rate of growth of this series over the 1940-1970 period is -0.005 W/m<sup>2</sup> per year, while — as mentioned in S4 — the corresponding rate of growth of RFGHG before the slowdown in RFCO<sub>2</sub> was 0.010 W/m<sup>2</sup> per year, about twice the rate (in absolute value) of RAER\_SENS. This result provides further evidence on the importance of the lack of increase in RFGHG as a factor contributing to this cooling period.

The second break in RAER\_SENS occurs prior to the recent slowdown in the warming, however the sign of the change in slope is positive, contributing to an increase in the aggregated radiative forcing. The reduction in the slope of RAER\_SENS is about 70%, going from -0.0061 W/m<sup>2</sup> per year to -0.002 W/m<sup>2</sup> per year after the break.

The sensitivity analysis shows that our findings are robust to the use of an alternative series for the radiative forcing of the direct effects of tropospheric aerosols.

## **S8. Relation to previous results in the literature.**

In this section, a brief review of some results from attribution studies based on time-series models is presented and compared to the findings presented in this paper.

Two recent papers<sup>48,49</sup> are related to the present study, in the sense that they use a similar regression-based approach to filter the effects of AMO from global temperatures in order to extract a clearer warming signal from the data. Nevertheless, the estimates and interpretation of the warming trend and of its nonlinearities are very different to those presented here. After establishing a close relation between the low-frequency oscillation in the Central England Temperature record and the AMO, these papers show that the oscillatory pattern associated

with the AMO has been occurring before the industrial times, a result supported by other previous publications<sup>3,44</sup>. This natural multidecadal oscillation, superimposed on a secular warming trend, is proposed for explaining the non-monotonic warming shown by global temperatures during the 20th century, including the cooling of the mid-century and the recent slowdown in the warming. The authors argue that the warming trend has remained almost constant since 1910 at 0.07–0.08 °C/decade<sup>48</sup>. The most significant differences between the results from these papers and those presented here can be attributed to the use of a linear trend to represent the warming trend, as well as to the lack of testing for breaks in the trend function. The present paper shows that once a common secular trend is identified between the aggregate radiative forcing and temperature series, the changes in the radiative forcing of the greenhouse gases are uncovered as relevant factors contributing to both the pause of the warming of the mid-20th century and the slowdown in the warming experienced since the 1990s. These results hold even when the AMO, and other forcings (e.g., aerosols), are taken into account. Assuming a linear trend as a proxy for the warming trend has the important disadvantage of potentially masking nonlinearities in the trend function as part of the realizations of the noise component.

The existence of a slowdown in global temperatures in the last decades has been recently analyzed in the literature<sup>50</sup>. After adjustments to remove the effects of factors such as El Niño/Southern Oscillation, volcanic aerosols and solar variability on the short-term variability of global and hemispheric temperatures, the authors find that the warming trend is constant over the period 1979-2010. The differences with the results presented in this paper can be explained by the following differences in the analyses. Our results about the slowdown are based on removing the low-frequency variability produced by AMO which, if it is not taken into account, can exaggerate the warming trend when in its positive phase<sup>1,2,3</sup>. In addition, the results presented here are based on more powerful approaches to test for breaks. Finally, our conclusions are supported by the analysis of the underlying radiative forcing trend, which clearly shows a decrease in its rate of growth since the mid 1990s.

A recent publication based on polynomial cointegration<sup>51</sup> has proposed that both solar irradiance and global temperature series are integrated of order one, while greenhouse gases and aerosols forcing are integrated of order two. These authors find that, although greenhouse gases and aerosols forcing cointegrate to an integration order of one, the stochastic trend in temperatures and in the aggregated anthropogenic forcing are independent, precluding the existence of a long-term relationship between them. The authors conclude that previous cointegration-based studies<sup>5,4,35</sup> have overlooked the differences in the order of integration of these series and accordingly delivered spurious results.

In light our findings, these conflicting results can easily be explained. Consider a simulation experiment consisting of 5,000 realizations of synthetic temperature and radiative forcing series with a sample spanning 1880 to 2010. The synthetic temperature and radiative forcing are constructed as trend stationary processes with a break in the slope of their trend functions with parameters as in the Supplementary Tables 2 and 9 for G and RFGHG, respectively. AR(1) stationary noise components are added to the broken trend functions with the first order autoregressive coefficients and variances estimated from the residuals obtained from detrending G (NASA) and RFGHG using the corresponding broken trend. As shown in Supplementary Table 18, when applying a standard ADF test to these series in levels, the unit root null is rejected only 34.54% of the times for the synthetic G series and never for the corresponding synthetic RFGHG, even though by construction these time series are trend stationary with a break. These results would erroneously lead to the argument that these

series are at least integrated of order one. Applying the ADF test to the first-differenced synthetic series to test for a possible integration of order two, the results are very similar to those described in some cointegration-based attribution studies<sup>5,51</sup>: temperatures are found to be integrated of order one while radiative forcing series are integrated of order two. These results are spurious and, as shown in the literature, they may lead to contrasting cointegration results and inferences depending on the order of integration selected.

This type of spurious results when testing for the integration order of a series that exhibits structural breaks has been explained in the econometrics literature<sup>11,52</sup>. A change in the slope of a trending series will completely dominate the noise component and induce a non-rejection of the unit root<sup>11</sup>. In consequence the series will be classified as integrated at least of order one. In a second step, a unit root test is applied to the first-differenced series. The broken trend process in first-differences becomes a process with a change in mean. According to Perron<sup>52</sup>, testing for unit roots when there is a break in the mean also increases the probability of a non-rejection of the null hypothesis but the result depends heavily on the properties of the noise component. For the radiative forcing series, the amplitude of the noise is very small in comparison to the signal, producing a clear bias in the sum of the autoregressive coefficients towards unity, hence a non-rejection. For temperatures, the noise is large in comparison to the signal and thus the bias produced by the change in the level is too small to induce a non-rejection. Consequently, this leads to the erroneous finding of temperatures and radiative forcing series being integrated of order one and two, respectively.

In light of these results, our model can be said to be encompassing in the sense that it not only provides a better description of the data but can also explain the results from other studies.

### **S9. Break dates and changes in measurement of CO<sub>2</sub> and other radiative forcing factors.**

A potential problem when conducting observational-based attribution studies relates to how the radiative forcing time series are constructed. Most of the forcing series are a combination of in situ measurements over the last decades and reconstructions from ice core and firn data. In the case of atmospheric CO<sub>2</sub>, CH<sub>4</sub> and N<sub>2</sub>O instrumental measurement started in 1958, 1981 and 1978 respectively, while concentrations prior to these dates are based on analyses of ice cores data<sup>25,53</sup>. In addition, CFCs amounts before 1978 are inferred from industrial production data.

There may be a concern that the results presented in this paper could be sensitive to the fact that there were changes in the statistical properties of RFCO<sub>2</sub> in 1958, which is a dominant component of RFGHG and TRF. This is, however, not the case unless the changes induce a change in the slope of the trend of RFCO<sub>2</sub>, which as argued below is unlikely. Measurement errors will not affect any of the results about the nature of the underlying trend function.

Changes in anthropogenic emissions of CO<sub>2</sub> do not occur independently from those of other greenhouse gases as they are in general a byproduct of economic activities and growth. As discussed previously in the literature, the growth rates of atmospheric GHG accelerated after WWII<sup>53</sup>. The components of RFGHG show a break around the mid 20th century: 1964, 1943, 1966, 1956 and 1957 for the radiative forcing of N<sub>2</sub>O, CH<sub>4</sub>, CO<sub>2</sub>, CFC<sub>11</sub> and CFC<sub>12</sub>, respectively. Note that in all cases the dates of the start of instrumental observations are outside the 95% confidence intervals of the estimated break dates (Supplementary Table 19). Consequently, the results from analyzing the individual trend of the components of RFGHG

support the existence of a common underlying driver and are not the product of changes in measurement.

## S10. Supplementary Methods.

A Review<sup>54</sup> contains a description of the various methods used in this paper.

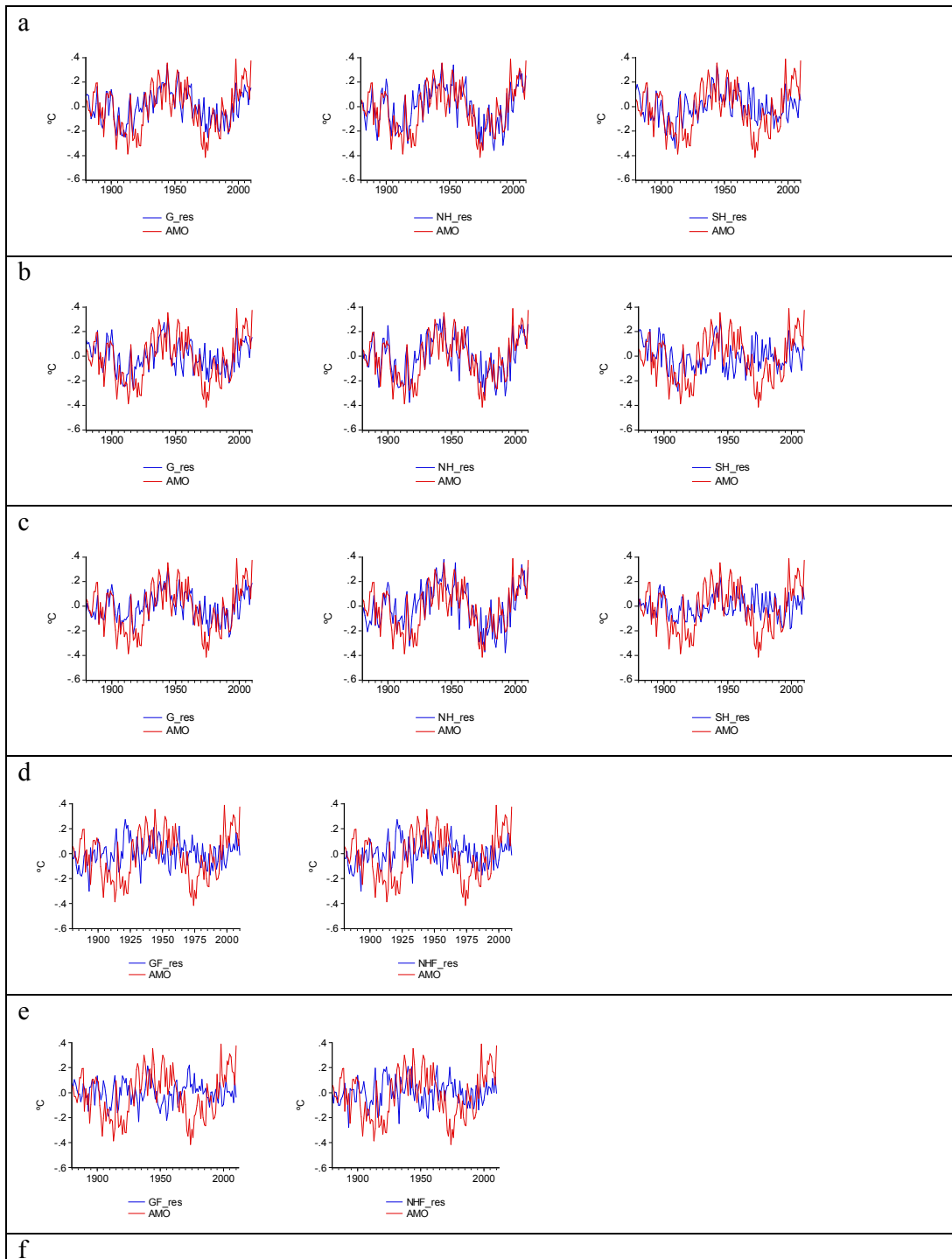
## Supplementary Information References

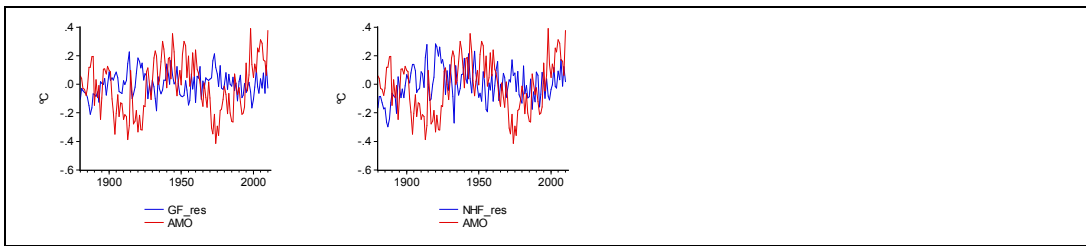
1. Wu, Z., Huang, N. E., Wallace, J. M., Smoliak, B. V. & Chen, X. On the time-varying trend in global-mean surface temperature. *Clim. Dyn.* **37**, 759–773 (2011).
2. Swanson, K. L., Sugihara, G. & Tsonis, A. A. Long-term natural variability and the 20th century climate change. *Proc. Natl. Acad. Sci.* **106**, 16120-16123 (2009).
3. Knudsen, M. F., Seidenkrantz, M. S., Jacobsen, B. H. & Kuijpers, A. Tracking the Atlantic Multidecadal Oscillation through the last 8,000 years. *Nat. Comm.* **2**, 178 (2011).
4. Kaufmann, R. K. & Stern, D.I. Evidence for human influence on climate from hemispheric temperature relations. *Nature* **388**, 39–44 (1997).
5. Kaufmann, R.K., Kauppi, H. & Stock, J.H. Emissions, concentrations, & temperature: a time series analysis. *Clim. Change* **77**, 249-278.
6. Gay, C., Estrada, F. & Sánchez, A. Global and hemispheric temperature revisited. *Clim. Change* **94**, 333–349 (2009).
7. Bierens, H.J. Nonparametric nonlinear cotrending analysis, with an application to interest and inflation in the United States. *JBES* **18**, 323-337 (2000).
8. Dickey, D.A. & Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Statist. Assoc.* **74**, 427–431 (1979).
9. Said, E. & Dickey, D. A. Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika* **71**, 599–607 (1984).
10. Elliott, G., Rothenberg, T.J. & Stock, J.H. Efficient tests for an autoregressive unit root. *Econometrica* **64**, 813–836 (1996).
11. Perron, P. The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* **57**, 1361–1401 (1989).
12. Estrada, F., Perron, P., Gay, C. & Martínez, B. A time series analysis of the 20th century climate simulations produced for the IPCC's AR4. *PLoS ONE* **8(3)**, e60017 (2013).
13. Estrada, F. & Perron, P. Breaks, trends and the attribution of climate change: a time-series analysis. Working paper, Department of Economics, Boston University. Available at <http://people.bu.edu/perron/workingpapers.html>.
14. Perron, P. & Yabu, T. Testing for shifts in trend with an integrated or stationary noise component. *JBES* **27**, 369-396 (2009).
15. Perron, P. Further evidence on breaking trend functions in macroeconomic variables. *J. Econom.* **80**, 355-385 (1997).
16. Kim, D. & Perron, P. Unit root tests allowing for a break in the trend function under both the null and the alternative hypotheses. *J. Econom.* **148**, 1-13 (2009).
17. Perron, P. & Vogelsang, T.J. Erratum: The great crash, the oil price shock and the unit root hypothesis. *Econometrica* **61**, 248-249 (1993).
18. Estrada, F., Gay, C. & Sánchez, A. Reply to “Does temperature contain a stochastic trend? Evaluating conflicting results by Kaufmann et al”. *Clim. Change* **101**, 407-414.
19. Elliott, G. On the robustness of cointegration methods when regressors almost have unit roots. *Econometrica* **66**, 149-158 (1998).
20. Gonzalo, J. & Lee T.H. Pitfalls in testing for long run relationships. *J. Econom.* **86**, 129-154 (1998).

21. Leybourne, S.J. & Newbold, P. Spurious rejections by cointegration tests induced by structural breaks. *Applied Economics* **35**, 1117-1121 (2003).
22. Perron, P. & Zhu, X. Structural breaks with deterministic and stochastic trends, *J. Econom.* **129**, 65-119 (2005).
23. Zwiers, F.W. & Weaver, A.J. The causes of 20th century warming. *Science* **290**, 5499 (2000).
24. Peixoto, J. P. & Oort, A. H. *Physics of climate*. American Institute of Physics, New York, USA, 520 pp.
25. IPCC-WGI: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp (2007).
26. Engle, R.F. & Granger, C.W.J. Co-integration and error correction: Representation, estimation and testing. *Econometrica* **55**, 251-276 (1987).
27. Engle, R.F. & Kozicki, S. Testing for common features. *JBES* **11**, 369-395 (1993).
28. Hansen, J. & Lebedeff, S. Global trends of measured surface air temperature. *J. Geophys. Res.* **92**, 13 345–13 372 (1987).
29. Jones, P.D., Raper, S.C.B., Bradley, R.S., Diaz, H.F., Kelly, P.M. & Wigley, T.M.L. Northern Hemisphere surface air temperature variations: 1851-1984. *J. Clim. App. Meteorol.* **25**, 161-179 (1986).
30. Jones, P.D., Raper, S.C.B. & Wigley, T.M.L. Southern Hemisphere surface air temperature variations: 1851-1984. *J. Clim. App. Meteorol.* **25**, 1213-1230 (1986).
31. Jones, P.D., Wigley, T.M.L. & Wright, P.B. Global temperature variations between 1861 and 1984. *Nature* **322**, 430-434 (1986).
32. Thompson, D.W.J., Kennedy, J.J., Wallace, J.M. & Jones, P.D. A large discontinuity in the mid-twentieth century in observed global mean surface temperature. *Nature* **453**, 646–649 (2008).
33. Hansen, J., Ruedy, R. & Sato, M. Global Temperature in 2011, Trends, and Prospects. <http://data.giss.nasa.gov/gistemp/2011/> 18 January 2012.
34. Muller, R.A., Curry, J., Groom, D., Jacobsen, R., Perlmutter, S., Rohde, R., Rosenfeld, A., Wickham, C. & Wurtele, J. Decadal variations in the global atmospheric land temperatures. Available at: <http://berkeleyearth.org/pdf/berkeley-earth-decadal-variations.pdf>.
35. Kaufmann, R.K., Kauppi, H., Mann, M.L. & Stock, J.H. Reconciling anthropogenic climate change with observed temperature 1998–2008. *Proc. Natl. Acad. Sci.* **108**, 11790-11793 (2011).
36. Kejriwal, M. & Perron, P. A sequential procedure to determine the number of breaks in trend with an integrated or stationary noise component. *J. Time Ser. Anal.* **31**, 305–328 (2010).
37. Hansen, J.E. & Sato, M. Trends of measured climate forcing agents. *Proc. Natl. Acad. Sci.* **98**, 14778-14783 (2001).
38. Velders, G.J.M., et al. The importance of the Montreal Protocol in protecting climate. *Proc. Natl. Acad. Sci.* **104**, 4814-19 (2007).
39. Kai, F.U., Tyler, S.C., Randerson, J.T. & Blake, D.R. Reduced methane growth rate explained by decreased Northern Hemisphere microbial sources. *Nature* **476**, 194–197 (2011).
40. Schwartz, S.E. Determination of Earth's transient and equilibrium climate sensitivities from observations over the twentieth century: strong dependence on assumed forcing. *Surv. Geophys.* **33**, 745–777 (2012).

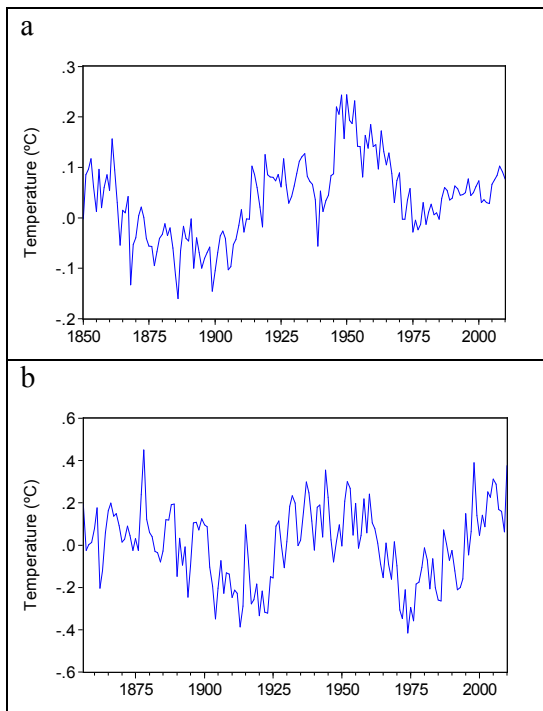
41. Gregory, J.M. & Forster, P.M. Transient climate response estimated from radiative forcing and observed temperature change. *J. Geophys. Res.* **113**, D23105 (2008).
42. Gregory, J.M. Vertical heat transports in the ocean and their effect on time-dependent climate change. *Clim. Dyn.* **16**, 501–515 (2000).
43. Held, I.M., Winton, M., Takahashi, K., Delworth, T., Zeng, F. & Vallis, G.K. Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J. Clim.* **23**, 2418–2427 (2010).
44. Knight, J.R., Allan, R.J., Folland, C.K. & Vellinga, M. A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophys. Res. Lett.* **32**, L20708 (2005).
45. Schwartz, S.E. Uncertainty requirements in radiative forcing of climate change. *J. Air & Waste Manage. Assoc.* **54**, 1351–1359 (2004).
46. Haywood, J.M., & Boucher, O. Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review. *Rev. Geophys.* **38**, 513-543 (2000).
47. Meinshausen, M., Smith, S.J., Calvin, K.V., Daniel, J.S., Kainuma, M.L.T., Lamarque, J.F., Matsumoto, K., Montzka, S.A., Raper, S.C.B., Riahi, K., Thomson, A.M., Velders, G.J.M. & van Vuuren, D. The RCP greenhouse gas concentrations and their extension from 1765 to 2300. *Clim. Change* **109**, 213-241 (2011).
48. Tung, K.K. & Zhou, J. Using data to attribute episodes of warming and cooling in instrumental records. *Proc. Natl. Acad. Sci.* **110**, 2058-2063 (2013)
49. Zhou J. & Tung, K.K. Deducing multidecadal anthropogenic global warming trends using multiple regression analysis. *J. Atmos. Sci.* **70**, 3-8 (2013).
50. Foster, G. & Rahmstorf, S. Global temperature evolution 1979-2010. *Environ. Res. Lett.* **6**, 044022 (2011).
51. Beenstock, M., Reingewertz, Y. & Paldor, N. Polynomial cointegration tests of anthropogenic impact on global warming. *ESD* **3**, 173-188 (2012).
52. Perron, P. Testing for a unit root in a time series with a changing mean. *JBES* **8**, 153-162 (1990).
53. Hansen, J. & Sato, M. Greenhouse gas growth rates. *Proc. Natl. Acad. Sci.* **101**, 16109-16114 (2004).
54. Estrada F. & Perron P. Detection and attribution of climate change through econometric methods. Working paper, Department of Economics, Boston University. Available at <http://people.bu.edu/perron/workingpapers.html>.

## Supplementary Figures and Legends

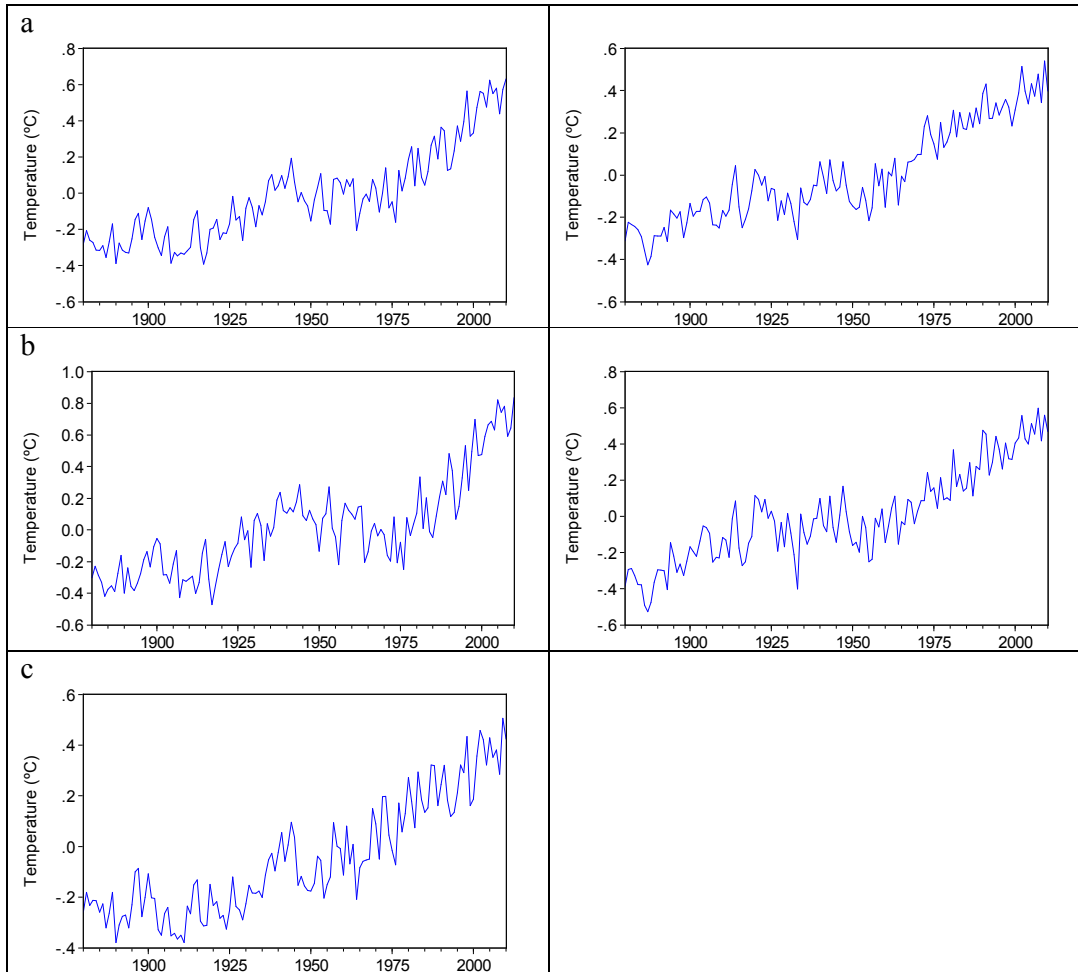




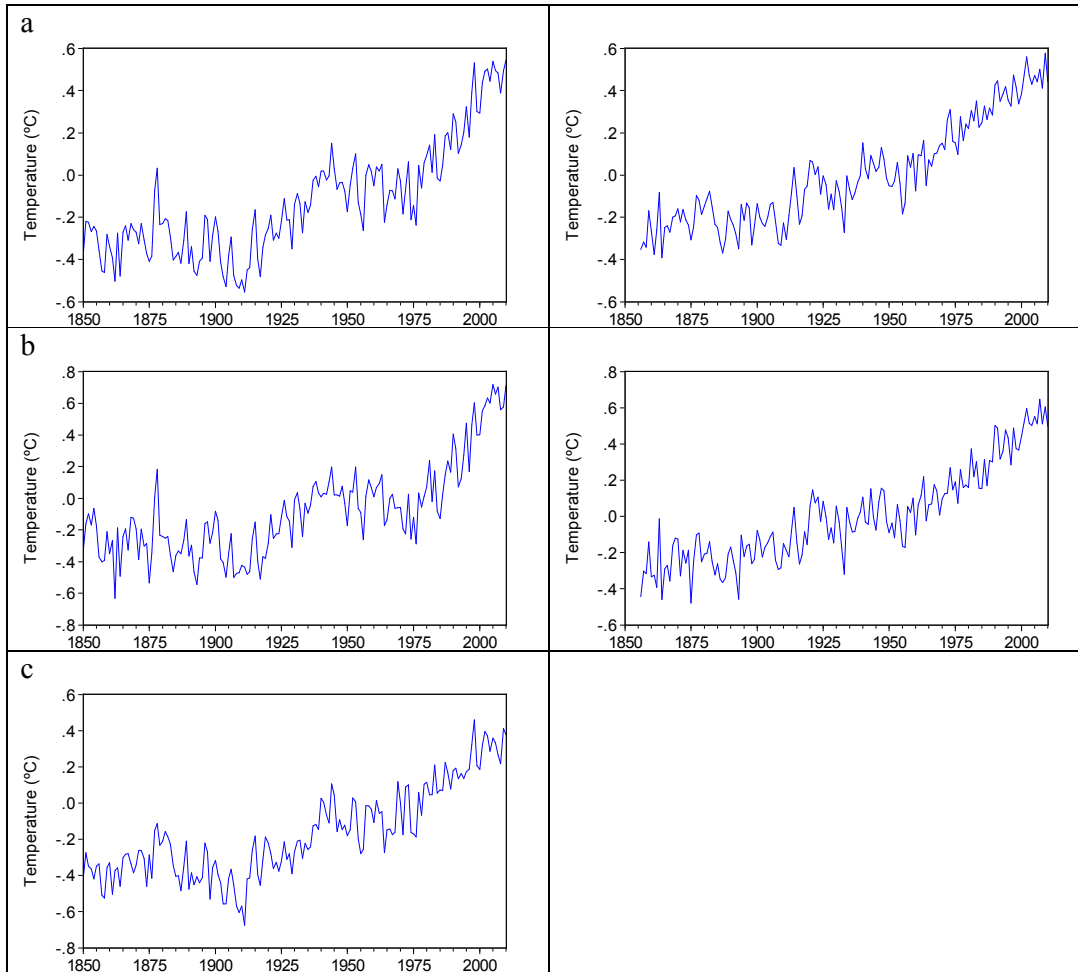
Supplementary Figure 1. Panel a) shows the detrended G, NH and SH series ( $G_{res}$ ,  $NH_{res}$  and  $SH_{res}$ , respectively) from the HadCRUT4 dataset in comparison with AMO. Panel b) shows the detrended G, NH and SH series ( $G_{res}$ ,  $NH_{res}$  and  $SH_{res}$ , respectively) from the NASA dataset in comparison with AMO. Panel c) shows the detrended G, NH and SH series ( $G_{res}$ ,  $NH_{res}$  and  $SH_{res}$ , respectively) from the HadCRUT3 dataset in comparison with AMO. Panel d) shows the detrended filtered G and NH ( $GF_{res}$ ,  $NHF_{res}$ , respectively) from the HadCRUT4 dataset in comparison with AMO. Panel e) shows the detrended filtered G and NH ( $GF_{res}$ ,  $NHF_{res}$ , respectively) from the HadCRUT3 dataset in comparison with AMO. Panel f) shows the detrended filtered G and NH ( $GF_{res}$ ,  $NHF_{res}$ , respectively) from the NASA dataset in comparison with AMO.



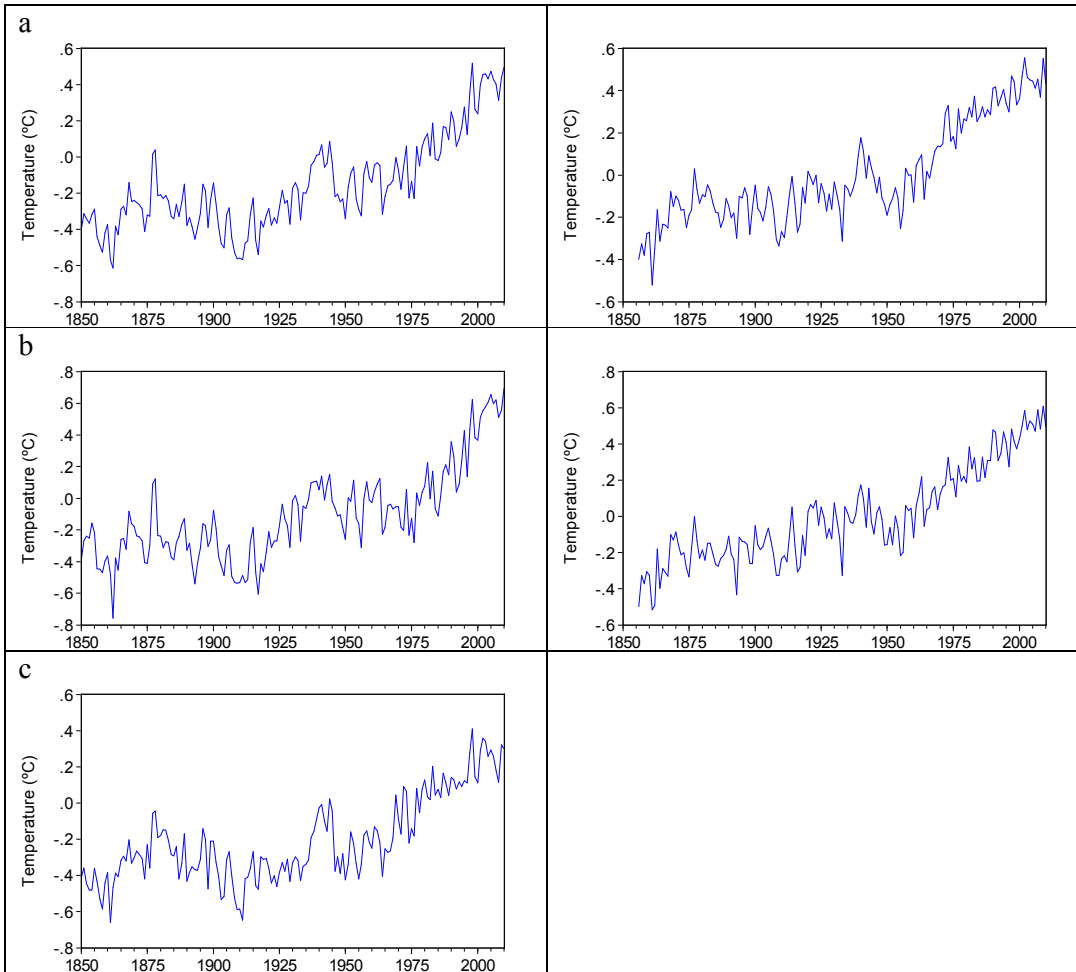
Supplementary Figure 2. Panel a) shows the difference between SH from HadCRUT4 and HadCRUT3. Panel b) shows AMO.



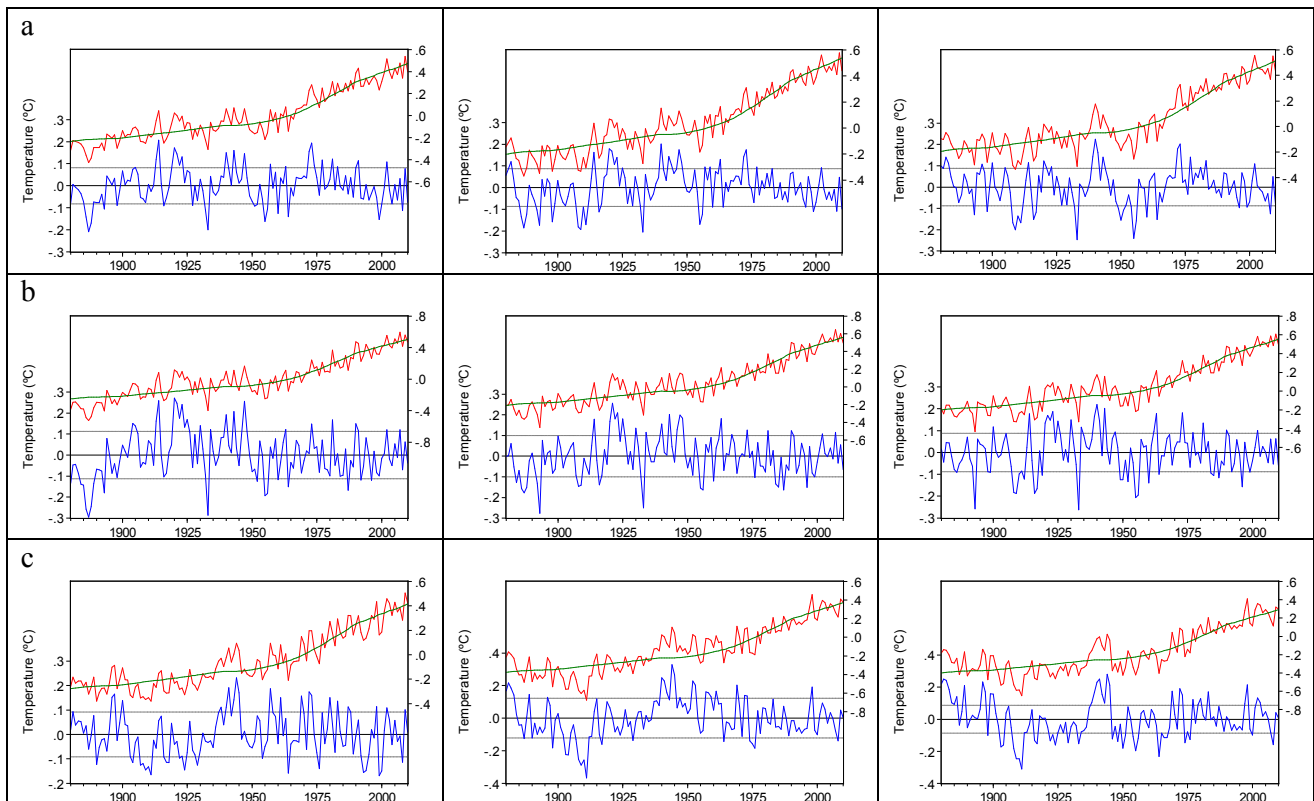
Supplementary Figure 3. Panels a) and b) show the unfiltered (left) and filtered (right) G and NH temperature series. Panel c) shows the unfiltered SH temperatures. NASA dataset.



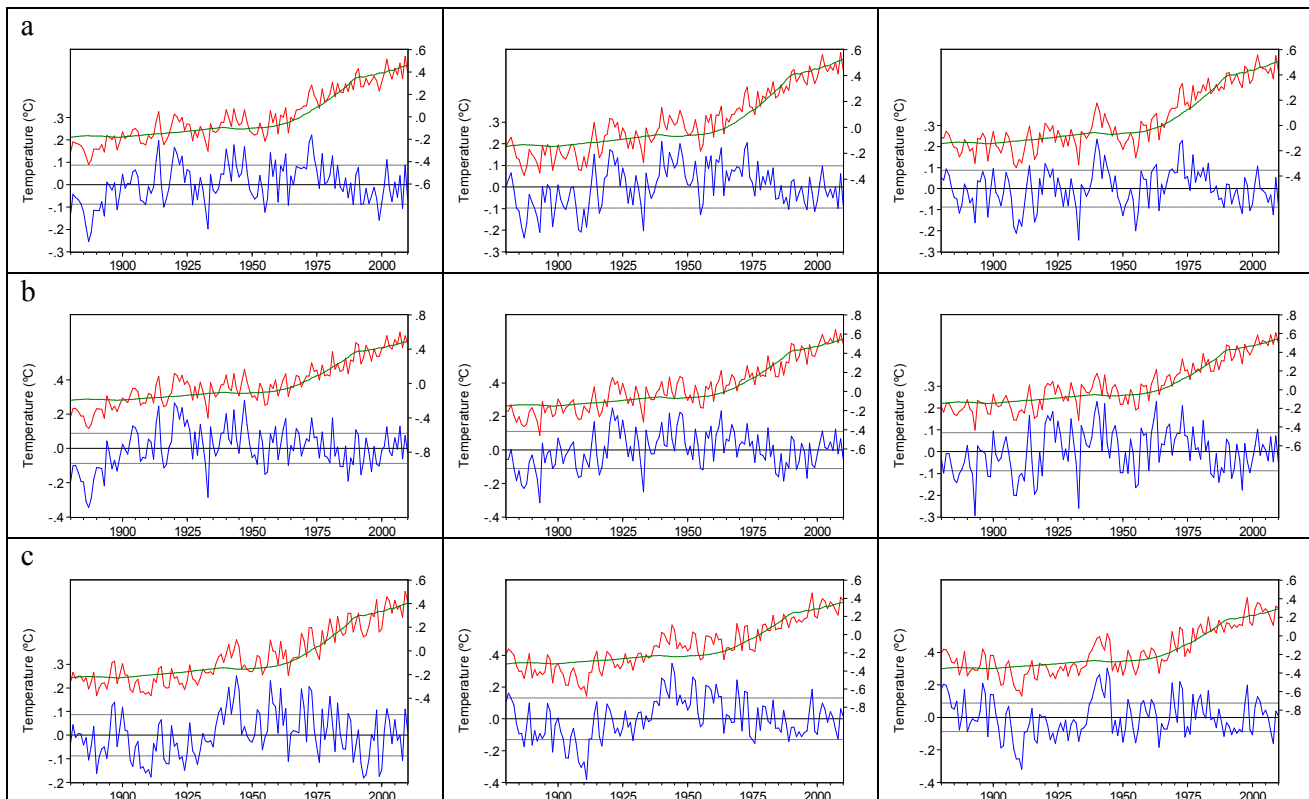
Supplementary Figure 4. Panels a) and b) show the unfiltered (left) and filtered (right) G and NH temperature series. Panel c) shows the unfiltered SH temperatures. HadCRUT4 dataset.



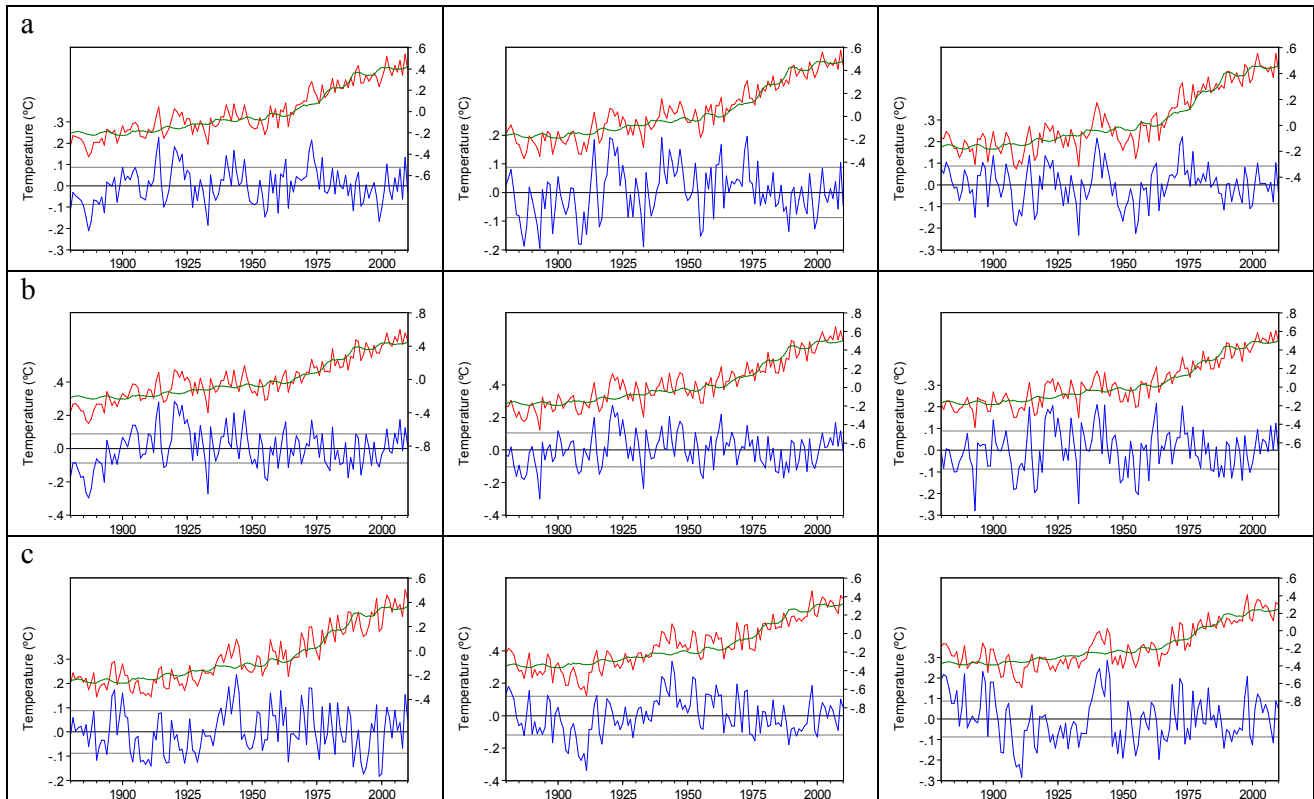
Supplementary Figure 5. Panels a) and b) show the unfiltered (left) and filtered (right) G and NH temperature series. Panel c) shows the unfiltered SH temperatures. HadCRUT3 dataset.



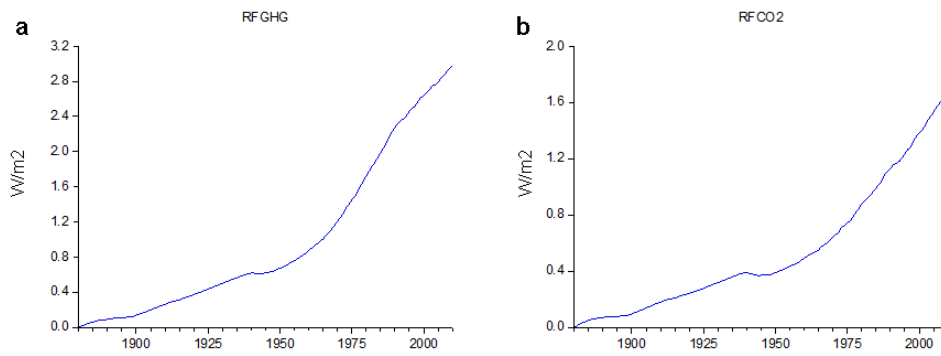
Supplementary Figure 6a. Panels a), b) and c) show the observed G, NH (filtered) and SH (unfiltered) temperature series, the fitted temperature series (units on the right axis) using RFGHG and the corresponding temperature residuals (units on the left axis) for G, NH and SH, respectively. NASA (left), HadCRUT4 (center) and HadCRUT3 (right) datasets.



Supplementary Figure 6b. Panels a), b) and c) show the observed G, NH (filtered) and SH (unfiltered) temperature series, the fitted temperature series (units on the right axis) using TRF\* and the corresponding temperature residuals (units on the left axis) for G, NH and SH, respectively. NASA (left), HadCRUT4 (center) and HadCRUT3 (right) datasets.

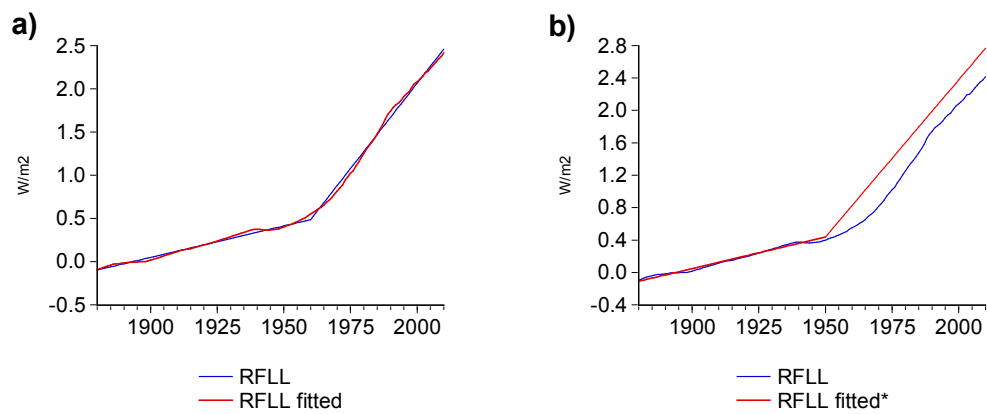


Supplementary Figure 6c. Panels a), b) and c) show the observed G, NH (filtered) and SH (unfiltered) temperature series, the fitted temperature series (units on the right axis) using TRF and the corresponding temperature residuals (units on the left axis) for G, NH and SH, respectively. NASA (left), HadCRUT4 (center) and HadCRUT3 (right) datasets.

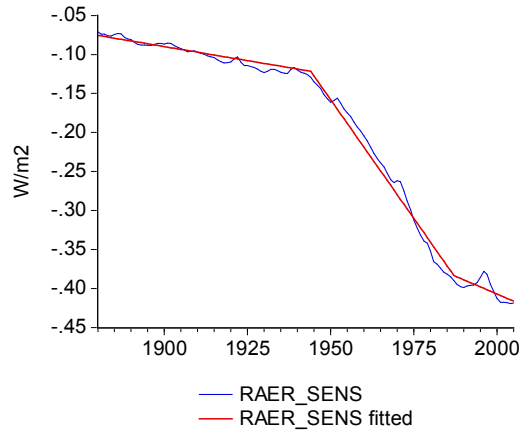


lementary Figure 7. Panel a) RFGHG. Panel b) RFCO2.

Supp



Supplementary Figure 8. Panel a) shows RFL and the fitted piece-wise linear trend with a break occurring in 1960. Panel b) shows a counterfactual experiment by fitting a piece-wise linear trend in which the slowdown did not occur (RFL fitted\*).



Supplementary Figure 9. Time series plot of RAER\_SENS and the fitted trend function with two breaks: 1944 and 1987.

## Supplementary Tables

**Supplementary Table 1. Perron-Yabu structural change test applied to unfiltered temperature series.**

Series	Exp Wald statistic
$G^{C4}$	<b>9.55 [11.15]</b>
$NH^{C4}$	<b>13.59</b>
$SH^{C4}$	<b>15.22 [2.01]</b>
$G^{C3}$	<b>8.91</b>
$NH^{C3}$	<b>8.02</b>
$SH^{C3}$	<b>10.74</b>
$G^N$	<b>13.96</b>
$NH^N$	<b>10.87</b>
$SH^N$	<b>7.56 [4.05]</b>

C4, C3 and N superscripts indicate HadCRUT4, HadCRUT3 and NASA datasets, respectively. The Exp Wald statistic associated with the second break is given in brackets. Bold and italic figures denote statistical significance at the 1% and 5% levels, respectively.

**Supplementary Table 2. Tests for a unit root allowing for a one-time break in the trend function applied to unfiltered temperature series.**

Series	$T_b$	k	$\hat{\mu}$	$t_{\hat{\mu}}$	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\gamma}$	$t_{\hat{\gamma}}$	$\hat{\alpha}$	$t_{\hat{\alpha}}$	$S(\hat{e})$	$t_{\alpha}(\hat{\lambda}_r^{AO})$
G(S1) <sup>C4</sup>	1910	0	-0.30	<b>-10.62</b>	-0.0007	-0.95	0.0064	<b>5.51</b>	0.54	-5.91 <sup>a</sup>	0.1188	-6.08 <sup>a</sup>
G(S2) <sup>C4</sup>	1976	0	-0.17	<b>-6.72</b>	0.0026	<b>3.71</b>	0.0134	<b>7.97</b>	0.35	-6.63 <sup>a</sup>	0.1009	-5.99 <sup>a</sup>
NH <sup>C4</sup>	1982	0	-0.36	<b>-14.10</b>	0.0028	<b>8.68</b>	0.0219	<b>10.21</b>	0.49	-7.36 <sup>a</sup>	0.1526	-6.37 <sup>a</sup>
SH(S1) <sup>C4</sup>	1909	0	-0.33	<b>-12.22</b>	-0.0010	-1.44	0.0069	<b>6.24</b>	0.49	-6.49 <sup>a</sup>	0.1137	-5.14 <sup>a</sup>
SH(S2) <sup>C4</sup>	1976	0	-0.74	<b>-13.03</b>	0.0060	<b>10.40</b>	0.0042	<b>2.63</b>	0.40	-6.74 <sup>a</sup>	0.1071	-6.00 <sup>a</sup>
G <sup>C3</sup>	1971	0	-0.40	<b>-18.09</b>	0.0022	<b>7.45</b>	0.0136	<b>10.17</b>	0.59	-6.37 <sup>a</sup>	0.1259	-5.61 <sup>a</sup>
NH <sup>C3</sup>	1984	0	-0.40	<b>-15.37</b>	0.0030	<b>9.27</b>	0.0243	<b>9.45</b>	0.58	-6.46 <sup>a</sup>	0.1544	-5.91 <sup>a</sup>
SH <sup>C3</sup>	1955	0	-0.39	<b>-17.03</b>	0.0011	<b>3.13</b>	0.0099	<b>10.26</b>	0.55	-6.86 <sup>a</sup>	0.1221	-6.26 <sup>a</sup>
G <sup>N</sup>	1978	0	-0.32	<b>-16.58</b>	0.0039	<b>12.15</b>	0.0129	<b>9.04</b>	0.40	-7.45 <sup>a</sup>	0.0909	-6.40 <sup>a</sup>
NH <sup>N</sup>	1985	0	-0.31	<b>-11.81</b>	0.0042	<b>10.02</b>	0.0235	<b>9.00</b>	0.46	-6.90 <sup>a</sup>	0.1238	-5.57 <sup>a</sup>
SH(S3) <sup>N</sup>	1910	0	-0.21	<b>-8.38</b>	-0.0034	<b>-2.77</b>	0.0112	<b>5.56</b>	0.43	-4.87 <sup>a</sup>	0.0840	-4.17 <sup>b</sup>
SH(S4) <sup>N</sup>	1954	0	-0.05	-1.32	0.0032	-1.32	0.0125	<b>4.39</b>	0.16	-7.50 <sup>a</sup>	0.0909	-7.03 <sup>a</sup>

C4, C3 and N superscripts indicate HadCRUT4, HadCRUT3 and NASA datasets, respectively. (S1), (S2), (S3) and (S4) indicate that the tests are applied using the

samples 1850-1976, 1920-2012, 1880-1944 and 1935-2010, respectively. The regression model for the unit root tests is  $\tilde{y}_t = \alpha y_{t-1} + \sum_{i=1}^k a_i \Delta \tilde{y}_{t-i} + e_t$  where  $\tilde{y}_t$  is

the detrended series, i.e. the residuals from equation (1) in the main text estimated by OLS using the stated break date. The symbols used are defined as follows:  $T_b$  is the estimate of the break date; k is the number of lagged differences added to correct for serial autocorrelation;  $S(\hat{e})$  is the standard error of the regression;  $\hat{\mu}$ ,  $\hat{\beta}$ ,  $\hat{\gamma}$  are the regression coefficients of the trend function in equation (1) of the main text and  $t_{\hat{\mu}}$ ,  $t_{\hat{\beta}}$ ,  $t_{\hat{\gamma}}$  the corresponding t-statistic values. Bold numbers denote statistical significance at the 5% level.  $\hat{\alpha}$  is the estimate of the sum of the autoregressive coefficients and  $t_{\hat{\alpha}}$  is the Perron<sup>15</sup> unit root test statistic.  $t_{\alpha}(\hat{\lambda}_r^{AO})$  is the Kim-Perron<sup>16</sup> unit root test statistic. a, b, c denote statistical significance at the 1%, 5% and 10%, respectively.

**Supplementary Table 3. Break date estimates and 95% confidence intervals for the unfiltered temperature series.**

Series	Break date
G <sup>C4</sup>	1976 (1969, 1983)
G <sup>C4*</sup>	1910 (1889, 1931)
NH <sup>C4</sup>	1982 (1974, 1990)
SH <sup>C4</sup>	1976 (1962, 1990)
SH <sup>C4*</sup>	1909 (1894, 1924)
G <sup>C3</sup>	1971 (1960, 1982)
NH <sup>C3</sup>	1984 (1976, 1992)
SH <sup>C3</sup>	1955 (1943, 1967)
G <sup>N</sup>	1978 (1971, 1985)
NH <sup>N</sup>	1985 (1978, 1992)
SH <sup>N</sup>	1954 (1947, 1961)
SH <sup>N*</sup>	1910 (1902, 1918)

C4, C3 and N superscripts indicate HadCRUT4, HadCRUT3 and NASA datasets, respectively. \*superscript indicates the results are obtained using the break dates in the early 20th century. 95% confidence intervals are shown in parenthesis. Long-run variance estimated using the Bartlett kernel with Andrews' bandwidth selection method.

**Supplementary Table 4. Test for nonlinear co-trending amongst unfiltered G, NH, SH, RFGHG and TRF.**

r	HadCRUT4	HadCRUT3	NASA	10% critical region	5% critical region
1	0.03	0.03	0.03	>0.12	>0.15
2	0.06	0.06	0.06	>0.17	>0.20
3	0.09	0.09	0.07	>0.21	>0.25
4	0.18	0.14	0.15	>0.25	>0.30
5	<b>0.43</b>	<b>0.38</b>	<b>0.38</b>	>0.28	>0.33

Bold figures indicate statistical significance at the 10% level.

**Supplementary Table 5. Correlation coefficients between AMO, NAO, PDO, SOI and the residuals from regressing the unfiltered temperatures series on TRF.**

	AMO	NAO	PDO	SOI
$G^{C4}$ _res	<b>0.75</b>	-0.17	0.05	<b>-0.21</b>
$NH^{C4}$ _res	<b>0.79</b>	-0.15	0.02	-0.11
$SH^{C4}$ _res	<b>0.54</b>	-0.17	0.09	<b>-0.31</b>
$G^{C3}$ _res	<b>0.71</b>	<b>-0.24</b>	<b>0.22</b>	<b>-0.29</b>
$NH^{C3}$ _res	<b>0.79</b>	<b>-0.19</b>	0.12	-0.16
$SH^{C3}$ _res	<b>0.37</b>	<b>-0.24</b>	<b>0.31</b>	<b>-0.40</b>
$G^N$ _res	<b>0.72</b>	-0.16	0.11	<b>-0.19</b>
$NH^N$ _res	<b>0.75</b>	-0.12	0.06	-0.09
$SH^N$ _res	<b>0.41</b>	<b>-0.18</b>	0.17	<b>-0.29</b>

C4, C3 and N superscripts indicate HadCRUT4, HadCRUT3 and NASA datasets, respectively. Bold figures indicate statistical significance at the 5% level.

**Supplementary Table 6. Standard unit root tests applied to AMO.**

	ADF	DF-GLS	ERS-PO
AMO	<b>-3.31</b> (2)	<b>-2.45</b> (2)	<b>2.16</b> (2)

The model specification includes a constant. The lag length and bandwidth shown in parenthesis. Figures in bold indicate that the statistic is significant at the 5% level. For the ADF and DF-GLS tests the lag length was selected via the Akaike Information Criterion. For the ERS-PO, the autoregressive spectral density estimator is used with the lag length selected via the Akaike Information Criterion.

**Supplementary Table 7. Nonlinear co-trending test applied to unfiltered and filtered G, NH.**

r	G HadCRUT4	NH HadCRUT4	G HadCRUT3	NH HadCRUT3	G NASA	NH NASA	10% critical region	5% critical region
1	0.08	0.08	0.08	0.08	0.08	0.08	>0.12	>0.15
2	<b>0.28</b>	<b>0.21</b>	<b>0.28</b>	<b>0.25</b>	<b>0.28</b>	<b>0.21</b>	>0.17	>0.20

Bold figures indicate statistical significance at the 10% level.

**Supplementary Table 8. Perron-Yabu structural change test applied to filtered temperature and radiative forcing series.**

Series	Exp Wald F-statistic
$G^{C4F}$	<b>17.06</b>
$NH^{C4F}$	<b>22.92</b>
$SH^{C4}$	<i>2.01</i>
$G^{CF}$	<b>15.16</b>
$NH^{CF}$	<b>16.11</b>
$SH^C$	<b>10.74</b>
$G^{NF}$	<b>12.34</b>
$NH^{NF}$	<b>6.25</b>
$SH^N$	<b>11.95</b>
RFGHG	<b>6.74</b>
TRF*	<b>3.24</b>
TRF	<b>7.92</b>

C4, C3 and N superscripts indicate HadCRUT4, HadCRUT3 and NASA datasets, respectively. F superscript denotes filtered temperature series. Bold and italic figures denote statistical significance at the 1% and 5% levels, respectively.

**Supplementary Table 9. Tests for a unit root allowing for a one-time break in the trend function applied to filtered G and NH and unfiltered SH temperature series and RFGHG, TRF\* and TRF.**

Series	$T_b$	k	$\hat{\mu}$	$t_{\hat{\mu}}$	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\gamma}$	$t_{\hat{\gamma}}$	$\hat{\alpha}$	$t_{\hat{\alpha}}$	$S(\hat{e})$	$t_{\alpha}(\hat{\lambda}_{tr}^{AO})$
G <sup>C4F</sup>	1956	0	-0.31	<b>-16.61</b>	0.0030	<b>11.10</b>	0.0067	<b>7.70</b>	0.37	-7.96 <sup>a</sup>	0.0849	-7.84 <sup>a</sup>
NH <sup>C4F</sup>	1966	0	-0.34	<b>-17.40</b>	0.0035	<b>13.14</b>	0.0083	<b>8.55</b>	0.16	-10.60 <sup>a</sup>	0.1002	-9.41 <sup>a</sup>
SH(S2) <sup>C4</sup>	1976	0	-0.74	<b>-13.03</b>	0.0060	<b>10.40</b>	0.0042	<b>2.63</b>	0.40	-6.74 <sup>a</sup>	0.1071	-6.00 <sup>a</sup>
G <sup>CF</sup>	1955	0	-0.26	<b>-13.50</b>	0.0021	<b>7.63</b>	0.0082	<b>10.97</b>	0.46	-7.57 <sup>a</sup>	0.0919	-7.16 <sup>a</sup>
NH <sup>CF</sup>	1959	0	-0.33	<b>-16.15</b>	0.0032	<b>11.09</b>	0.0075	<b>8.70</b>	0.30	-9.23 <sup>a</sup>	0.1002	-8.79 <sup>a</sup>
SH <sup>C</sup>	1955	0	-0.39	<b>-17.03</b>	0.0011	<b>3.13</b>	0.0099	<b>10.26</b>	0.55	-6.86 <sup>a</sup>	0.1221	-6.26 <sup>a</sup>
G <sup>NF</sup>	1956	0	-0.27	<b>-15.94</b>	0.0031	<b>9.01</b>	0.0063	<b>8.28</b>	0.34	-7.98 <sup>a</sup>	0.0788	-7.30 <sup>a</sup>
NH <sup>NF</sup>	1968	0	-0.31	<b>-14.09</b>	0.0039	<b>10.03</b>	0.0077	<b>6.31</b>	0.35	-7.84 <sup>a</sup>	0.1083	-7.33 <sup>a</sup>
SH(S4) <sup>N</sup>	1954	0	-0.05	-1.32	0.0032	-1.32	0.0125	<b>4.39</b>	0.16	-7.50 <sup>a</sup>	0.0909	-7.03 <sup>a</sup>
RFGHG	1960	7	-0.03	<b>-4.08</b>	0.0105	<b>64.95</b>	0.0351	<b>87.76</b>	0.90	-4.24 <sup>b</sup>	0.0399	-3.97 <sup>b</sup>
TRF*	1960	7	-0.03	<b>-2.67</b>	0.0036	<b>15.41</b>	0.0257	<b>41.92</b>	0.92	-3.78 <sup>c</sup>	0.0510	-3.99 <sup>b</sup>
TRF	1960	1	-0.09	<b>-5.34</b>	0.0064	<b>20.82</b>	0.0221	<b>28.98</b>	0.84	-4.58 <sup>a</sup>	0.0760	-4.11 <sup>b</sup>

C4, C3 and N superscripts indicate HadCRUT4, HadCRUT3 and NASA datasets, respectively. (S2) indicates that the test was applied using the sample 1920-2012. F

superscript denotes filtered temperature series. The regression model for the unit root tests is  $\tilde{y}_t = \alpha y_{t-1} + \sum_{i=1}^k a_i \Delta \tilde{y}_{t-i} + e_t$  where  $\tilde{y}_t$  is the detrended series, i.e. the residuals from equation (1) in the main text estimated by OLS using the stated break date. The symbols used are defined as follows:  $T_b$  is the estimate of the break date; k is the number of lagged differences added to correct for serial autocorrelation;  $S(\hat{e})$  is the standard error of the regression;  $\hat{\mu}$ ,  $\hat{\beta}$ ,  $\hat{\gamma}$  are the regression coefficients of the trend function in equation (1) of the main text and  $t_{\hat{\mu}}$ ,  $t_{\hat{\beta}}$ ,  $t_{\hat{\gamma}}$  the corresponding t-statistic values. Bold numbers denote statistical significance at the 5% levels.  $\hat{\alpha}$  is the estimate of the sum of the autoregressive coefficients and  $t_{\hat{\alpha}}$  is the Perron<sup>15</sup> unit root test statistic.  $t_{\alpha}(\hat{\lambda}_{tr}^{AO})$  is the Kim-Perron<sup>16</sup> unit root test statistic. a, b, c denote statistical significance at the 1%, 5% and 10%, respectively.

**Supplementary Table 10. Break date estimates and 95% confidence intervals for filtered G, NH, unfiltered SH, RFGHG, TRF\* and TRF.**

Series	Break date
G <sup>C4F</sup>	1956 (1946, 1966)
NH <sup>C4F</sup>	1966 (1958, 1974)
SH <sup>C4</sup>	1976 (1962, 1990)
G <sup>C3F</sup>	1955 (1945, 1965)
NH <sup>C3F</sup>	1959 (1950, 1968)
SH <sup>C3</sup>	1955 (1943, 1967)
G <sup>NF</sup>	1956 (1946, 1966)
NH <sup>NF</sup>	1968 (1956, 1980)
SH <sup>N</sup>	1955 (1947, 1961)
RFGHG	1960 (1959, 1961)
TRF*	1960 (1959, 1961)
TRF	1960 (1956, 1964)

C4, C3 and N superscripts indicate HadCRUT4, HadCRUT3 and NASA datasets, respectively. F superscript denotes filtered temperature series. 95% confidence intervals are shown in parenthesis. Long-run variance estimated using the Bartlett kernel with Andrews' bandwidth selection method.

**Supplementary Table 11. Test for nonlinear co-trending amongst filtered G, NH, unfiltered SH, RFGHG and TRF.**

r	HadCRUT4	HadCRUT3	NASA	10% critical region	5% critical region
1	0.03	0.03	0.03	>0.12	>0.15
2	0.06	0.07	0.07	>0.17	>0.20
3	0.09	0.08	0.08	>0.21	>0.25
4	0.15	0.15	0.17	>0.25	>0.30
5	<b>0.40</b>	<b>0.38</b>	<b>0.38</b>	>0.28	>0.33

Bold figures are significant at the 10% level.

**Supplementary Table 12. Kejriwal-Perron sequential structural change testing procedure applied to TRF, TRF\*, RFGHG and its components, RAER, RFAIE and RFBC.**

Series	Exp Wald statistic	Percent change	Break date
RFGHG	<b>5.00</b>	-25.61	1994 (1989, 1999)
TRF*	<b>6.15</b>	-50.66	1992 (1989, 1995)
TRF	<b>19.62</b>	-57.95	1991 (1988, 1994)
RFCO2	<b>2.43</b>	19.78	1996 (1993, 1999)
RFCH4	<b>4.79</b>	-73.35	1992 (1988, 1996)
RFCFC11	<b>3.01</b>	-115.47	1993 (1991, 1995)
RFCFC12	<b>4.54</b>	-92.85	1995 (1992, 1998)
N2O	0.51	--	--
RAER	0.14	--	--
RFAIE	<b>19.88</b>	32.48	1988 (1987, 1989)
RFBC	0.30	--	--

Bold figures are significant at the 1% level. Confidence intervals are given in parenthesis. Long-run variance estimated using the Bartlett kernel with Andrews' bandwidth selection method.

**Supplementary Table 13. Estimated transient climate sensitivities for filtered and unfiltered G.**

	$S_{tr}^{un}$	$S_{tr}^f$
$G^{C4}$	0.4269	0.3962
$G^{C3}$	0.4085	0.3793
$G^N$	0.3464	0.3872

C4, C3 and N superscripts indicate HadCRUT4, HadCRUT3 and NASA datasets, respectively.  $S_{tr}^{un}$ ,  $S_{tr}^f$  are the transient climate sensitivity estimated from equation (2) for the unfiltered and filtered global temperatures, respectively. All coefficients are significant at the 5% level.

**Supplementary Table 14. Test for a unit root allowing for a one-time break in the trend function applied to TRF\_SENS.**

Series	$T_b$	k	$\hat{\mu}$	$t_{\hat{\mu}}$	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\gamma}$	$t_{\hat{\gamma}}$	$\hat{\alpha}$	$t_{\hat{\alpha}}$	$S(\hat{e})$	$t_{\alpha}(\hat{\lambda}_{tr}^{AO})$
TRF_SENS	1962	2	-0.17	<b>-12.88</b>	0.0107	<b>43.32</b>	0.0350	<b>49.44</b>	0.77	-5.60 <sup>a</sup>	0.0622	-4.52 <sup>b</sup>

The regression model for the unit root tests is  $\tilde{y}_t = \alpha y_{t-1} + \sum_{i=1}^k a_i \Delta \tilde{y}_{t-i} + e_t$  where  $\tilde{y}_t$  is the detrended series, i.e. the residuals from equation (1) in the main text estimated by OLS using the stated break date. The symbols used are defined as follows:  $T_b$  is the estimate of the break date; k is the number of lagged differences added to correct for serial autocorrelation;  $S(\hat{e})$  is the standard error of the regression;  $\hat{\mu}$ ,  $\hat{\beta}$ ,  $\hat{\gamma}$  are the regression coefficients of the trend function in equation (1) of the main text and  $t_{\hat{\mu}}$ ,  $t_{\hat{\beta}}$ ,  $t_{\hat{\gamma}}$  the corresponding t-statistic values. Bold numbers denote statistical significance at the 5% level.  $\hat{\alpha}$  is the estimate of the sum of the autoregressive coefficients and  $t_{\hat{\alpha}}$  is the Perron<sup>15</sup> unit root test statistic.  $t_{\alpha}(\hat{\lambda}_{tr}^{AO})$  is the Kim-Perron<sup>16</sup> unit root test statistic. a, b, c denote statistical significance at the 1%, 5% and 10%, respectively.

**Supplementary Table 15. Test for nonlinear co-trending amongst unfiltered G, NH, SH, RFGHG and TRF\_SENS.**

r	HadCRUT4	HadCRUT3	NASA	10% critical region	5% critical region
1	0.04	0.04	0.03	>0.12	>0.15
2	0.06	0.05	0.04	>0.17	>0.20
3	0.10	0.09	0.09	>0.21	>0.25
4	0.18	0.16	0.19	>0.25	>0.30
5	<b>0.38</b>	<b>0.37</b>	<b>0.35</b>	>0.28	>0.33

Bold figures indicate statistical significance at the 10% level.

**Supplementary Table 16. Test for nonlinear co-trending amongst filtered G, NH, unfiltered SH, RFGHG and TRF\_SENS.**

r	HadCRUT4	HadCRUT3	NASA	10% critical region	5% critical region
1	0.04	0.04	0.03	>0.12	>0.15
2	0.06	0.06	0.06	>0.17	>0.20
3	0.09	0.09	0.09	>0.21	>0.25
4	0.16	0.16	0.19	>0.25	>0.30
5	<b>0.35</b>	<b>0.35</b>	<b>0.35</b>	>0.28	>0.33

Bold figures indicate statistical significance at the 10% level.

**Supplementary Table 17. Kejriwal-Perron sequential structural change testing procedure applied to RAER\_SENS.**

Series	Exp Wald statistic	Percent change	Break date
RAER_SENS (1st break)	<b>2.59</b>	646.59	1944 (1941, 1947)
RAER_SENS (2nd break)	<i>1.64</i>	-69.99	1987 (1983, 1991)

Bold and italic figures are significant at the 1%, and 10% levels, respectively. 95% confidence intervals are given in parenthesis. Long-run variance estimated using the Bartlett kernel with Andrews' bandwidth selection method.

**Supplementary Table 18. Percentage of rejections of the null hypothesis for the simulated G and RFGHG series.**

	G sim	RFGHG sim
Levels	34.54%	0.00%
First-differences	99.78%	25.02%

Rejection rate of a standard ADF test based on simulations with 5,000 replications. The model specification includes a constant and a trend. Lag length selected by the Akaike Information Criterion.

**Supplementary Table 19. Dates of the start of instrumental observations and estimates of the break dates with 95% confidence intervals for the break dates in the components of RFGHG during the mid 20th century.**

Series	Date of start of instrumental observation	Break date	Exp Wald statistic
RFCO2	1958	1966 (1965, 1967)	<i>1.35</i>
RFCH4	1981	1943 (1942, 1946)	<b>8.10</b>
RFN2O	1978	1964 (1962, 1966)	<b>111.73</b>
RFCFC11	1978	1956 (1955, 1957)	<b>165.82</b>
RFCFC12	1978	1957 (1956, 1958)	<b>171.14</b>

95% confidence intervals are shown in parenthesis. Long-run variance estimated using the Bartlett kernel with Andrews' bandwidth selection method. Bold and italic figures denote statistical significance at 1% and 10% levels.

# Statistically-derived contributions of diverse human influences to 20th century temperature changes

Francisco Estrada<sup>1,2</sup>, Pierre Perron<sup>3</sup> & Benjamín Martínez-López<sup>1</sup>

<sup>1</sup>Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Ciudad Universitaria, Circuito Exterior, 04510 Mexico, DF, Mexico, <sup>2</sup>Institute for Environmental Studies, Vrije Universiteit, Amsterdam, Netherlands, <sup>3</sup>Department of Economics, Boston University, 270 Bay State Rd. Boston, MA, 02215, USA.

## Supplementary Methods

### 1 Perron-Yabu testing procedure for structural changes in the trend function.

Perron<sup>10</sup> showed that the presence of structural changes can have considerable implications when investigating time-series properties by means of unit root tests. This creates a circular problem given that most of the tests for structural breaks require to correctly identify if the data generating process is stationary or integrated. Depending on whether the process is stationary or integrated the limit distribution of these tests are different and, if the process is misidentified, the tests will have poor properties.

The Perron-Yabu<sup>11</sup> test was designed explicitly to address the problem of testing for structural changes in the trend function of a univariate time series without any prior knowledge as to whether the noise component is stationary or contains an autoregressive unit root. The approach of Perron-Yabu builds on previous work of the same authors who analyzed the problem of hypothesis testing on the slope coefficient of a linear trend model when no information about the nature,  $I(0)$  or  $I(1)$ , of the noise component is available<sup>12</sup>.

We present the case of a model with a one-time structural break in the slope of the trend function with an autoregressive noise component of order one (AR(1)). A more detailed presentation of this case and of other structural change models and extensions can be found in the original Perron-Yabu article. Consider the following data generating process:

$$\begin{aligned}y_t &= x_t' \Psi + u_t \\ u_t &= \alpha u_{t-1} + e_t\end{aligned}\tag{1}$$

for  $t = 1, \dots, T$ ,  $e_t \sim i.i.d. (0, \sigma^2)$ ,  $x_t$  is a  $(r \times 1)$  vector of deterministic components, and  $\Psi$  is a  $(r \times 1)$  vector of unknown parameters which are model specific and described in the next paragraphs. The initial condition  $u_0$  is assumed to be bounded in probability. The autoregressive coefficient is such that  $|\alpha| \leq 1$  and therefore, both integrated and stationary

errors are allowed. The interest is testing the null hypothesis  $R\Psi = \gamma$  where  $R$  is a  $(q \times r)$  full rank matrix and  $\gamma$  is a  $(q \times 1)$  vector, where  $q$  is the number of restrictions. The restrictions are used to test for the presence of a structural change in the trend function. For this purpose, Perron-Yabu consider three models where a change in intercept and/or slope in the trend function occurs. In what follows, the break date is denoted  $T_1 = [\lambda T]$  for some  $\lambda \in (0, 1)$ , where  $[\cdot]$  denotes the largest integer that is less than or equal to the argument and  $1(\cdot)$  is the indicator function.

The model to test for a one-time change in the slope of the trend function is specified with  $x_t = (1, t, DT_t)'$  and  $\Psi = (\mu_0, \beta_0, \beta_1)'$  where  $DT_t = 1(t > T_1)(t - T_1)$  so that the trend function is joined at the time of the break. The hypothesis of interest is  $\beta_1 = 0$ . The testing procedure is based on a Quasi Feasible Generalized Least Squares approach that uses a superefficient estimate of the sum of the autoregressive parameters  $\alpha$  when  $\alpha = 1$ . The estimate of  $\alpha$  is the OLS estimate obtained from an autoregression applied to detrended data and is truncated to take a value 1 when the estimate is in a  $T^{-\delta}$  neighborhood of 1. This makes the estimate "super-efficient" when  $\alpha = 1$  and implies that in the case of a known break date, inference on the slope parameter can be performed using the standard Normal or Chi-square distribution whether  $\alpha = 1$  or  $|\alpha| < 1$ . Theoretical arguments and simulation evidence show that  $\delta = 1/2$  is the appropriate choice. When the break date is unknown, the limit distribution is nearly the same in the I(0) and I(1) cases when considering the Exp functional of the Wald test across all permissible dates for a specified equation. Hence, it is possible to have tests with nearly the same size in both cases. To improve the finite sample properties of the test, they also use a bias-corrected version of the OLS estimate of  $\alpha$  as suggested by Roy and Fuller<sup>13</sup>. The testing procedure suggested by the authors is:

1. For any given break date, detrend the data by Ordinary Least Squares (OLS) to obtain the residuals  $\hat{u}_t$ ;
2. Estimate an AR(1) model for  $\hat{u}_t$  yielding the estimate  $\hat{\alpha}$ ;
3. Use  $\hat{\alpha}$  to get the Roy-Fuller biased corrected estimate  $\hat{\alpha}_M$ ;
4. Apply the truncation

$$\hat{\alpha}_{MS} = \begin{cases} \hat{\alpha}_M & \text{if } |\hat{\alpha}_M - 1| > T^{-\frac{1}{2}} \\ 1 & \text{if } |\hat{\alpha}_M - 1| \leq T^{-\frac{1}{2}} \end{cases}$$

5. Apply a Generalized Least Squares (GLS) procedure with  $\hat{\alpha}_{MS}$  to obtain the estimates of the coefficients of the trend and the variance of the residuals and construct the standard Wald-statistic  $W_{FMS}$ ;
6. Since the break date is assumed to be unknown, the 5 steps above must be repeated for all permissible break dates to construct the Exp functional of the Wald test denoted by

$$Exp-W_{FS} = \log \left[ T^{-1} \sum_{\Lambda} \exp \left( \frac{1}{2} W_{FMS}(\lambda) \right) \right]$$

where  $\Lambda = \{\lambda; \epsilon \leq \lambda \leq 1 - \epsilon\}$  for some  $\epsilon > 0$ . Commonly used values are  $\epsilon = 0.05$  and  $\epsilon = 0.15$ .

## 2 Kerjiwal-Perron sequential procedure to determine the number of breaks in trend with an integrated or stationary noise component.

Kerjiwal-Perron<sup>14</sup> considered the problem of selecting the number of breaks in the trend function of a univariate time series without any prior knowledge about whether the noise component is stationary or contains an autoregressive unit root. Their test is an extension of the Perron-Yabu<sup>11</sup> procedure that allows for testing the null hypothesis of  $l$  changes against the alternative hypothesis of  $(l + 1)$  changes.

Following the data generating process described in (1), the model to test for the presence of  $(l + 1)$  breaks in the slope of the trend function is specified as  $x_t = (1, t, DT_{1t}, \dots, DT_{(l+1)t})'$  and  $\Psi = (\mu_0, \beta_0, \beta_1, \dots, \beta_{l+1})'$  where  $DT_{it} = 1 (t > T_i^0) (t - T_i^0)$  so the trend function is joined at the time of the break and where  $T_i^0$  denotes the true  $i$ th break date. The interest is in testing the null hypothesis of  $l$  breaks in the slope coefficient (i.e.,  $\beta_{l+1} = 0$ ) against the alternative hypothesis of  $(l + 1)$  breaks. We again present an overview of the case of an AR(1) noise process. A complete description of this case and of its extension to a more general structure for the error term  $u_t$  can be found in the original article of Kerjiwal-Perron<sup>14</sup>. The implementation of the sequential test for the null hypothesis of  $l$  breaks against the alternative of  $(l + 1)$  breaks is as follows. The first step consists in estimating the break dates  $\hat{T}_1, \dots, \hat{T}_l$  as global minimizers of the sum of squared residuals (SSR) from the model with  $l$  breaks estimated by OLS:

$$(\hat{T}_1, \dots, \hat{T}_l) = \arg \min_{(\hat{t}_1, \dots, \hat{t}_l)} SSR(\hat{T}_1, \dots, \hat{T}_l)$$

This can be achieved efficiently using the algorithm of Bai and Perron<sup>15</sup>. One then tests for the presence of an additional break in each of the  $(l + 1)$  segments using the estimated

partition  $(\hat{T}_1, \dots, \hat{T}_l)$ . The following regression estimated by OLS is used to construct the test over the  $i$ th segment ( $i = 1, \dots, l + 1$ ):

$$y_t = x_t^{(i)'} \Psi^{(i)} + u_t, \text{ for } t = \hat{T}_{i-1} + 1, \dots, \hat{T}_i \quad (2)$$

where  $x_t^{(i)} = (1, t - \hat{T}_{i-1}, (t - k) I(t > k))'$ ,  $\Psi^{(i)} = (\mu_0^{(i)}, \beta_0^{(i)}, \beta_1^{(i)})'$ ,  $k = [T\tau]$ ,  $\tau \in \Lambda_{i,\epsilon} = \{\tau : \hat{\lambda}_{i-1} + (\hat{\lambda}_i - \hat{\lambda}_{i-1})\epsilon \leq \tau \leq \hat{\lambda}_i + (\hat{\lambda}_i - \hat{\lambda}_{i-1})\epsilon\}$ , with  $\hat{\lambda}_i = \hat{T}_i/T$  and  $T_0 = 0$ . We again set  $\epsilon = 0.15$ .

It is important to note that the trend included in the  $i$ th segment is  $(t - \hat{T}_{i-1})$  instead of  $t$ . This permits ensuring that the initial conditions are the same across segments. The residuals  $\hat{u}_t^{(i)}$  from regression (2) are then used to compute the OLS estimate of  $\alpha$ . This OLS estimate in turn is then used to construct the superefficient estimate of  $\alpha$  denoted as  $\alpha_s^{(i)}$ . As in the Perron-Yabu procedure, the feasible GLS approach is used for obtaining the exponential functional of the Wald statistic over all permissible break dates. Once the  $Exp - W_{FMS}^{(i)}$  is obtained for  $i = 1, \dots, l + 1$ , the sequential test is defined by

$$F_T(l + 1 | l) = \max_{1 \leq i \leq l+1} \left\{ Exp - W_{FMS}^{(i)} \right\}.$$

The decision rule is to conclude in favour of a model with  $(l + 1)$  breaks if the maximum of the  $Exp - W_{FMS}^{(i)}$  tests is sufficiently large. Kerjiwal-Perron showed that in both I(0) and I(1) cases, the asymptotic critical values for the sequential test can be obtained from the relevant quantiles of the limit distribution for the single break test.

### 3 Perron and Kim-Perron unit root tests with a one-time break in the trend function

As shown in Perron<sup>10</sup>, the sum of the first order autoregressive coefficients is highly biased towards unity if there is a shift in the trend function. In this case, the unit root null is hardly rejected even if the series is composed of *i.i.d.* disturbances around a trend. Furthermore, if the break occurs in the slope of the trend function, unit root tests are not consistent, i.e., the null hypothesis of a unit root cannot be rejected even asymptotically.

Perron<sup>10</sup> proposed an extension of the Augmented Dickey-Fuller (ADF) test<sup>16,17</sup> that allows for a one-time break in the trend function of a univariate time series. Three different model specifications were considered: the ‘‘crash’’ model that allows for an exogenous change in the level of the series; the ‘‘changing growth’’ model that permits an exogenous change in the rate of growth; and a third model that allows both changes. For this test, the break

dates are treated as exogenous in the sense of intervention analysis<sup>18</sup>, separating what can and cannot be explained by the noise in a time series. Our interest centers in the “changing growth” model, which can be briefly described as follows. The null hypothesis is:

$$y_t = \mu_1 + y_{t-1} + (\mu_2 - \mu_1) DU_t + e_t$$

where  $DU_t = 1$  if  $t > T_B$ , 0 otherwise;  $T_B$  refers to the time of the break, and  $A(L)e_t = B(L)v_t$ ,  $v_t \sim i.i.d. (0, \sigma^2)$ , with  $A(L)$  and  $B(L)$   $p$ th and  $q$ th order polynomials, respectively, in the lag operator. The innovation series  $\{e_t\}$  are  $ARMA(p, q)$  type with possibly unknown  $p, q$  orders. The alternative hypothesis is:

$$y_t = \mu_1 + \beta_1 t + (\beta_2 - \beta_1) DT_t^* + e_t$$

where  $DT_t^* = t - T_B$  if  $t > T_B$  and 0 otherwise. The “changing growth” model takes an “additive outlier” approach in which the change is assumed to occur rapidly and the regression strategy consists in first detrending the series according the following regression:

$$y_t = \mu + \beta_1 t + \beta_2 DT_t^* + \tilde{y}_t \quad (3)$$

Then an ADF regression is estimated on the residuals  $\tilde{y}_t$  as follows:

$$\tilde{y}_t = \alpha \tilde{y}_{t-1} + \sum_{i=1}^k c_i \Delta \tilde{y}_{t-i} + e_t \quad (4)$$

where the  $k$  lagged values of  $\Delta \tilde{y}_{t-i}$  are added as a parametric correction for autocorrelation. In the Perron<sup>10</sup> test the break is assumed to occur at a known date. Later, Perron<sup>19</sup> generalized the test for the case when the date of the break is unknown and he proposed determining the break point endogenously from the data. This is done estimating the break date by minimizing the sum of squared residuals from regression (3). The resulting unit root test is then the t-statistic for testing that  $\alpha = 1$  in regression (4) estimated by OLS. The critical values of the limit distribution of the test are tabulated in Perron<sup>19</sup>.

A problem with most procedures for testing for unit roots in the presence of a one-time break that occurs at an unknown date is that the change in the trend function is allowed only under the alternative hypothesis of a stationary noise component<sup>19,20,21</sup>. As consequence, it is possible that a rejection occurs when the noise is  $I(1)$  and there is a large change in the slope of the trend function. A method that avoids this problem is that of Kim-Perron<sup>22</sup>. Their procedure is based on a pre-test for a change in the trend function, namely the Perron-Yabu test<sup>11</sup>. If this pre-test rejects, the limit distribution of the unit root test is then the same as

if the break date was known<sup>10,23</sup>. This is very advantageous since when a break is present the test has much greater power. It was also shown in simulations to maintain good size in finite samples and that it offers improvements over other commonly used methods. The testing procedure under the additive outlier approach for the changing growth model consists in the following steps:

1. Obtain an estimate of the break date  $\hat{T}_B = \hat{\lambda}T$  by minimizing the sum of squared residuals using regression (3). Then construct a window around that estimate defined by a lower bound  $T_l$  and an upper bound  $T_h$ . Depending on the sample size a window of 6-8 observations is commonly used. Note that, as shown by Kim and Perron<sup>22</sup>, the results are not sensitive to this choice;
2. Create a new data set  $\{y^n\}$  by removing the data from  $T_l + 1$  to  $T_h$ , and shifting down the data after the window by  $S(T) = y_{T_h} - y_{T_l}$ ; hence,

$$y^n = \begin{cases} y_t & \text{if } t \leq T_l \\ y_{t+t_h-t_l} - S(T) & \text{if } t > T_l \end{cases}$$

3. Perform the unit root test using the break date  $T_l$ . This is the  $t$ -test statistic for testing that  $\tilde{\alpha} = 1$  in the following regression estimated by OLS, denoted by  $t_\alpha(\hat{\lambda}_{tr}^{AO})$ :

$$\tilde{y}_t^n = \tilde{\alpha}\tilde{y}_t^n + \sum_{i=1}^k c_i \Delta \tilde{y}_{t-i}^n + \tilde{e}_t \quad (5)$$

where  $\hat{\lambda}_{tr} = T_l/T_r$ ,  $T_r = T - (T_h - T_l)$  and  $\tilde{y}_t^n$  is the detrended value of  $y^n$ .

The number of lags in (4) and (5) was chosen using the Schwarz Information Criterion (BIC). After choosing the lag length with the information criteria, Ljung-Box tests were performed on the residuals. In all cases, no evidence of remaining autocorrelation was found.

#### 4 Perron-Zhu methodology for estimating a confidence interval for the break date

Perron and Zhu<sup>24</sup> analyzed the consistency, rate of convergence and limiting distributions of parameter estimates in models where the trend exhibits a slope change at some unknown date and the noise component can be either stationary or have an autoregressive unit root. Their results are of particular relevance when considering the problem of selecting the break

date when testing for structural changes and for deriving the limiting distributions of unit root tests that allow for a one-time structural change that occurs at an unknown date, such as in the Kim-Perron test<sup>22</sup>. Another important practical application of deriving the limiting distribution of the estimate of the break date is that it permits forming a confidence interval for the break date.

Perron and Zhu<sup>24</sup> considered a total of six models with deterministic and stochastic trends. The random component was assumed to be either stationary or to contain a unit root, while for the deterministic component three cases were considered: 1) a first-order linear trend with a one-time change in the slope such that the trend function is joined at the time of the break; 2) a local disjoint broken trend; and 3) a global disjoint broken trend. Our case pertains to the first specification with a stationary noise component. The interested reader is referred to the original Perron and Zhu<sup>24</sup> paper for the specifications and limiting distributions for the other models considered by the authors. The deterministic part is specified as:

$$d_t = \mu_1 + \beta_1 t + \beta_b B_t,$$

where  $B_t$  is a dummy variable for the slope change defined by

$$B_t = \begin{cases} 0 & \text{if } t \leq T_1 \\ t - T_1 & \text{if } t > T_1 \end{cases}$$

with  $T_1 = \lambda T$  the break date and  $\lambda$  the break fraction. Note that at the time of the break, the slope coefficient changes from  $\beta_1$  to  $\beta_1 + \beta_b$  but that the trend function is continuous at  $T_1$ . This specification is therefore referred to as the “joint broken trend”. The estimation method is simply to select the break date that minimizes the sum of squared residuals from a regression of the series of interest  $y_t$  ( $t = 1, \dots, T$ ) on the regressors  $\{1, t, B_t\}$ , i.e., applying OLS to the model

$$y_t = \mu_1 + \beta_1 t + \beta_b B_t + u_t$$

Denote the resulting estimate by  $\hat{T}_1$  and the associated estimate of the break fraction by  $\hat{\lambda} = \hat{T}_1/T$ . They showed that the limit distribution of the break fraction  $\hat{\lambda}$  is:

$$T^{\frac{3}{2}}(\hat{\lambda} - \lambda) \rightarrow^d N \left( 0, \frac{4\sigma^2}{\left[ \lambda_0 (1 - \lambda_0) (\beta_b^0)^2 \right]} \right)$$

where  $\beta_b^0$  is the true value of the change in the slope parameter and  $\sigma^2 = \lim_{T \rightarrow \infty} E(\sum_{t=1}^T u_t)^2$  is the so-called long-run variance of  $u_t$ . Hence, it is straightforward to construct a confidence

interval using the estimates of  $\sigma^2$ ,  $\beta_b^0$  and  $\lambda_0$ . A common estimate of  $\sigma^2$  is based on a weighted sum of the autocovariance function of  $u_t$  of the form

$$\hat{S}_{w,T} = \hat{R}_v(0) + 2 \sum_{j=1}^{T-1} w(j, m) \hat{R}_u(j)$$

where  $\hat{R}_u(j) = T^{-1} \sum_{t=j+1}^T \hat{u}_t \hat{u}_{t-j}$  with  $\hat{u}_t$  the OLS residuals from regression (??). Here,  $w(j, m)$  is some weight function. A popular choice is the Bartlett triangular weight with  $w(j, m) = 1 - j/(m + 1)$  if  $j \leq m$  and 0 otherwise. Other choices are available such as the Parzen or Quadratic Spectral (e.g., Andrews, 1991). The parameter  $m$  is a bandwidth or truncation parameter. A popular method to select this parameter is due to Andrews (1991). We use the Bartlett weight function and Andrews' (1991) method to select  $m$ . This is done setting  $m = 1.1447(\alpha T)^{1/3}$ , where  $\alpha = 4\hat{\rho}^2/(1 - \hat{\rho}^2)^2$  with  $\hat{\rho}$  the OLS estimates of the coefficient from a first-order autoregression applied to  $\hat{u}_t$ .

## 5 Testing for a common long-run path

In this paper a simple approach is used to test for a common long-run path in the bivariate context. Assume two trend stationary variables of the form:

$$\begin{aligned} y_t &= d_t + u_t \\ x_t &= w_t + v_t \end{aligned}$$

where  $d_t$  and  $w_t$  are nonstationary components which may be composed of a wide class of linear and nonlinear trends, breaks in their slope parameters and/or shifts in their intercepts, and  $u_t$  and  $v_t$  are stationary noise components. The procedure is based on testing for remaining nonstationarities in the residuals of the following regression estimated by OLS:

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

If the individual nonstationary components  $d_t$  and  $w_t$  are present in  $y_t$  and  $x_t$  but not in  $\varepsilon_t$ , that is the residuals are found to be stationary, then it is said that  $y_t$  and  $x_t$  share the same long-run path. The existence of remaining non-common nonstationarities can be evaluated by applying standard unit root tests. This procedure is similar to the Engle-Granger two-steps cointegration test<sup>25</sup> but it does not require the assumption of unit roots in radiative forcing and temperature variables. Also, since all series are trend-stationary, the relevant critical values are those tabulated for standard unit root (or stationarity) tests with no deterministic terms included.

Since one of the principal issues investigated in this paper is the possibility of radiative forcing and temperature variables sharing common breaks in the slope of the trend function, we also applied a more specific test for evaluating it. Following the same rationale discussed above and the concept of common features of Engle-Kozicki<sup>26</sup>, the Perron-Yabu<sup>11</sup> test was used for evaluating if breaks that were previously detected in the series  $y_t$  and  $x_t$  were still present in the residuals  $\varepsilon_t$ .

The results from the Perron-Yabu test<sup>11</sup> and from the standard unit root tests applied to the residuals can provide sound evidence regarding the existence of a common long-term path, as well as of some associated common stylized facts such as breaks and transition periods.

## 6 Bierens nonparametric nonlinear co-trending test

Nonlinear co-trending is special case of the more general “common features” concept described by Engle and Kozicki<sup>26</sup>. The advantage of the test proposed by Bierens<sup>27</sup> is that the nonlinear trend does not have to be parameterized. The nonlinear trend stationarity model considered by Bierens can be expressed as follows:

$$z_t = g(t) + u_t$$

with

$$g(t) = \beta_0 + \beta_1 t + f(t)$$

where  $z_t$  is a  $k$ -variate time series,  $u_t$  is a  $k$ -variate zero-mean stationary process and  $f(t)$  is deterministic  $k$ -variate general nonlinear trend function that allows, in particular, structural changes. Nonlinear co-trending occurs when there exists a non-zero vector  $\theta$  such that  $\theta' f(t) = 0$ . Hence, the null hypothesis of this test is that the multivariate time series  $z_t$  is nonlinear co-trended, implying that there is one or more linear combinations of the time series that are stationary around a constant or a linear trend. Note that this test is a cointegration test in the case when it is applied to series that contain unit roots.

The nonparametric test for nonlinear co-trending is based on the generalized eigenvalues of the matrices  $M_1$  and  $M_2$  defined by:

$$M_1 = \frac{1}{n} \sum_{t=1}^n \hat{F} \left( \frac{t}{n} \right) \hat{F} \left( \frac{t}{n} \right)'$$

where

$$\hat{F}(x) = \frac{1}{n} \sum_{t=1}^{[nx]} (z_t - \hat{\beta}_0 - \hat{\beta}_1 t)$$

if  $x \in [n^{-1}, 1]$ ,  $\hat{F}(x) = 0$  if  $x \in [0, n^{-1})$  with  $\hat{\beta}_0$  and  $\hat{\beta}_1$  being the estimates of the vectors of intercepts and slope parameters in a regression of  $z_t$  on a constant and a time trend; and

$$M_2 = \left( \frac{1}{n} \right) \sum_{t=0}^{m-1} \left[ \left( \frac{1}{m} \right) \sum_{j=0}^{m-1} (z_{t-j} - \hat{\beta}_0 - \hat{\beta}_1 (t-j)) \right] \left[ \left( \frac{1}{m} \right) \sum_{j=0}^{m-1} (z_{t-j} - \hat{\beta}_0 - \hat{\beta}_1 (t-j)) \right]'$$

where  $m = n^\alpha$  with  $n$  equal to the number of observations and  $\alpha = 0.5$  as suggested by Bierens<sup>27</sup>. Solving  $|\hat{M}_1 - \lambda \hat{M}_2| = 0$  and denoting the solution  $\hat{\lambda}_r$ , the test statistic is  $n^{1-\alpha} \hat{\lambda}_r$ . The null hypothesis is that there are  $r$  co-trending vectors against the alternative of  $r - 1$  co-trending vectors. This test has a non-standard distribution and the critical values have been tabulated by Bierens<sup>27</sup>. The existence  $r$  co-trending vectors in  $r + 1$  series indicates the presence of  $r$  linear combinations of the series that are stationary around a linear trend and that these series share a single common nonlinear deterministic trend. Such a result indicate a strong secular co-movement in the  $r + 1$  series.

## Supplementary Methods References

1. Brohan, P., Kennedy, J.J., Harris, I., Tett, S.F.B. & Jones, P.D. Uncertainty estimates in regional and global temperature changes: a new dataset from 1850. *J. Geophys. Res.* **111**, D12106 (2006).
2. Hansen, J., Ruedy, R., Sato, M. & Lo, K. Global surface temperature change. *Rev. Geophys.* **48**, RG4004, doi:10.1029/2010RG000345 (2010).
3. Enfield, D.B., Mestas-Nunez, A.M. & Trimble, P.J. The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the continental U.S. *Geophys. Res. Lett.* **28**, 2077-2080 (2001).
4. Hansen, J., Sato, M., Kharecha, P. & von Schuckmann, K. Earth's energy imbalance and implications. arXiv, 1105.1140 (2011).
5. Hansen, J., Sato, M., Lacis, A., Ruedy, R., Tegen, I. & Mathews, E. Climate forcings in the industrial era. *Proc. Natl. Acad. Sci.* **95**, 12753-12758 (1998).
6. Hansen, J., Sato, M., Ruedy, R., Lacis, A. & Oinas, V. Global warming in the twenty-first century: An alternative scenario. *Proc. Natl. Acad. Sci.* **97**, 9875-9880 (2000).
7. Hansen, J., & Sato, M. Trends of measured climate forcing agents. *Proc. Natl. Acad. Sci.* **98**, 14778-14783 (2001).
8. Hansen, J., & Sato, M. Greenhouse gas growth rates. *Proc. Natl. Acad. Sci. USA* **101**, 16109-16114 (2004).
9. Boden, T.A., Marland, G. & Andres, R.J. Global, Regional, and National Fossil-Fuel CO<sub>2</sub> Emissions. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A. doi 10.3334/CDIAC/00001\_V2011 (2011).
10. Perron, P. The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* **57**, 1361-1401 (1989).
11. Perron, P. & Yabu, T. Testing for shifts in trend with an integrated or stationary noise component. *JBES* **27**, 369-396 (2009).
12. Perron, P. & Yabu, T. Estimating deterministic trends with an integrated or stationary noise component. *J. Econom.* **151**, 56-69 (2009).
13. Roy, A., & Fuller, W.A. Estimation for autoregressive processes with a root near one. *JBES* **19**, 482-493 (2001).
14. Kejriwal, M. & Perron, P. A sequential procedure to determine the number of breaks in trend with an integrated or stationary noise component. *J. Time Ser. Anal.* **31**, 305-328 (2010).

15. Bai, J. & Perron, P. Computation and analysis of multiple structural change models. *J. App. Econom.* **18**, 1-22 (2003).
16. Dickey D.A. & Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Statist. Assoc.* **74**, 427-431 (1979).
17. Said E. & Dickey, D.A. Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika* **71**, 599-607 (1984).
18. Box, G.E.P. & Tiao, G.C. Intervention analysis with applications to economic and environmental problems. *J. Amer. Statistical Assoc.* **70**, 70-79 (1975).
19. Perron, P. Further evidence on breaking trend functions in macroeconomic variables. *J. Econom.* **80**, 355-385 (1997).
20. Zivot, E. & Andrews, D.W.K. Further evidence on the Great Crash, the oil price shock, and the unit root hypothesis. *JBES* **10**, 251-270 (1992).
21. Vogelsang, T.J. & Perron, P. Additional tests for a unit root allowing the possibility of breaks in the trend function. *Int. Econ. Rev.* **39**, 1073-1100 (1998).
22. Kim, D. & Perron P. Unit root tests allowing for a break in the trend function under both the null and the alternative hypotheses. *J. Econom.* **148**, 1-13 (2009).
23. Perron, P. & Vogelsang, T.J. Erratum: The great crash, the oil price shock and the unit root hypothesis. *Econometrica* **61**, 248-249 (1993).
24. Perron, P. & Zhu, X. Structural breaks with deterministic and stochastic trends. *J. Econom.* **129**, 65-119 (2005).
25. Engle, R.F. & Granger, C.W.J. Co-integration and error correction: Representation, estimation and testing. *Econometrica* **55**, 251-276 (1987).
26. Engle, R.F. & Kozicki, S. Testing for common features. *JBES* **11**, 369-395 (1993).
27. Bierens, H.J. Nonparametric nonlinear cotrending analysis, with an application to interest and inflation in the United States. *JBES* **18**, 323-337 (2000).