

Extracting and analyzing the warming trend in global and hemispheric temperatures*

Francisco Estrada[†]

Universidad Nacional Autónoma de México and VU University Amsterdam

Pierre Perron[‡]

Boston University

May 26, 2016; Revised March 27, 2017

Abstract

This paper offers an updated and extended attribution analysis based on recently published versions of temperature and forcing datasets. It shows that both temperature and radiative forcing variables can be best represented as trend stationary processes with structural changes occurring in the slope of their trend functions and that they share a common secular trend and common breaks, largely determined by the anthropogenic radiative forcing. The common nonlinear trend is isolated and further evidence on the possible causes of the current slowdown in warming is presented. Our analysis offers interesting results in relation to the recent literature. Changes in the anthropogenic forcings are directly responsible for the hiatus as in Estrada et al. (2013a), while natural factors such as the Atlantic Multidecadal Oscillation, the Interdecadal Pacific Oscillation and the Southern Annular Mode, as well as the new temperature adjustments in Karl et al. (2105) contribute to weaken the signal. In other words, natural variability and data adjustments do not explain in any way the hiatus, they simply mask its presence.

Keywords: Climate change; warming hiatus; structural break; co-trending; principal component analysis.

*We thank the Editor-in-Chief, the Guest-Editor and two referees for useful comments.

[†]Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Ciudad Universitaria, Circuito Exterior, 0451 Mexico, DF, Mexico; and Institute for Environmental Studies, Vrije Universiteit, Amsterdam, Netherlands (feporrua@gmail.com).

[‡]Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215, USA (perron@bu.edu).

1 Introduction

The contributions of Working Groups I and II of the Intergovernmental Panel on Climate Change exhibit an ever-growing bulk of direct and indirect scientific evidence regarding the warming of the climate system during the last century and of the role anthropogenic activities (e.g., IPCC, 2013; IPCC, 2014). The detection and attribution of climate change is no longer limited to changes in climate variables. Studies have proposed that the warming signal is strong enough to affect other physical and biological systems to the extent that it can be tracked and attributed to human interference with the climate system as well (Zwiers and Hegerl, 2008; Rosenzweig et al., 2008; IPCC, 2014). An important part of the evidence on attribution is based on comparing observations to model predictions about what the state of a variety of systems would be (ranging from climate to natural and human systems) with or without anthropogenic changes in the atmosphere. One method that has shown to be of particularly importance for conducting attribution studies is the “optimal fingerprinting” (Hasselmann, 1979, 1997) which is based on a generalized multivariate regression for the detection and attribution of changes to externally forced climate change signals (IPCC, 2013). For this method, the dependent variable is usually an observed climate record and the covariates are composites of General Circulation Models output intended to represent the climate change signal. These “optimal detection analyses” that combine observed and modeled climate data have provided important evidence for supporting IPCC’s statements such as “most of the observed increase in global average temperatures since the mid-20th century is very likely due to the observed increase in anthropogenic greenhouse gas concentrations” (IPCC, 2013, see also Stott et al., 2006; Mitchell et al., 2001).

However, direct attribution of climate change to anthropogenic activities using observed climate and forcing variables remains challenging due the limited number of statistical methods available to investigate the existence of common long-term trends. The following is a brief background on the time-series based attribution literature, the interested reader is referred to Estrada and Perron (2014) for a literature review on this topic. Early examples of the application of modern statistical techniques to address the attribution of climate change go back to the 1990s with the work of Tol and de Vos (1993; 1998) and Stern and Kaufmann (2000), among others. The use of cointegration techniques (e.g., Engle and Granger, 1987) constituted a breakthrough on this subject and on how the possible presence of stochastic trends in temperature and forcing series could be interpreted. Nevertheless, although when these studies were published they benefited from some of the latest advances on the modeling of nonstationary time series, recently there have been important advances in testing for unit roots that have significantly changed econometric modeling (see Perron, 2006 for

a review) and that are useful to address the detection and attribution of climate change (Estrada et al., 2013a,b; Estrada and Perron, 2014; Gay et al., 2009; Estrada et al., 2010; Gil-Alana 2008). Recent publications have shown that the assumption on which these earlier attribution studies are based (i.e., temperatures and forcing variables being integrated processes) was not soundly tested and that there are strong reasons from both statistical and climate physics perspectives for questioning this assumption (Estrada et al., 2010; Gay et al., 2009; see also Triacca, 2001).

The attribution of climate change to human activities has been discussed at great length in the literature and, regardless of the differences in assumptions and methods (statistical- or physical-based), there is a general consensus about the existence of a common secular trend between temperatures and radiative forcing variables. Now the relevant questions to ask are related to the existence and causes of common features shown by the warming trend such as rapid warming, cooling periods, slowdowns and pauses. The study of particular features of this common long-term trend is very recent and has not yet been fully exploited (e.g., Estrada et al., 2013a). Learning from the existing common features in temperature and radiative forcing variables can help to better understand the drivers behind them and the impact climate policies can have. Recent developments in econometric modeling constitute a valuable set of tools to improve what is known about the observed warming (Estrada and Perron, 2014).

On the one hand, this paper offers an updated attribution analysis based on recently published versions of temperature and forcing datasets and state-of-the-art econometric techniques. It is shown that both temperature and radiative forcing variables can be best represented as trend stationary processes with structural changes occurring in the slope of their trend functions, the trend being defined as secular movement in time free of stochastic transitory fluctuations. The multivariate analysis strongly suggests that these variables share a common secular trend and common breaks, largely determined by the anthropogenic radiative forcing. On the other hand, it complements and extends the results in Estrada et al. (2013a) by means of a new approach based on a principal component analysis (PCA) to separate the common long-term trend imparted by radiative forcing from the natural variability component in global temperature series in order to tackle the debate on the existence of the recent slowdown in warming.

The PCA approach adopted in this paper helps to further investigate the potential causes of the reduced slowdown in the rate of increase of temperatures, the so-called hiatus. The main contenders to explain the so-called hiatus are : natural variability, data biases, changes in anthropogenic forcing (mostly CFC and methane reductions). Our results offer evidence against natural variability and data biases and in favor of changes in anthropogenic forcings. First, there is a clear reduction in the slope of the trend induced by anthropogenic forcings (associated with the first principal

component). This is the argument advanced by Estrada et al. (2013a) who established that the reduction was mostly due to reduced rate of increases in CFC and methane emissions. When this anthropogenic signal is mixed with the various other components, our results indicate that, in direct contrast to what has been advanced in the literature, the following contribute to blurring the signal and making the statistical evidence weaker: the Atlantic Multidecadal Oscillation (AMO), an oscillation likely related to the Interdecadal Pacific Oscillation (IPO) and to the Southern Annular Mode (SAM), as well as the new temperature adjustments in Karl et al. (2015). In other words, natural variability and data adjustments do not explain in any way the hiatus, they simply mask its presence. We believe these results are very interesting and should spur additional debates about the causes of the hiatus. The PCA results are also able to reconcile the difference in the estimates of the break dates across various series and across global and hemispheric temperatures.

The rest of the paper is structured as follows. Section 2 describes the data used. Section 3 presents the tests and procedures for the univariate analysis, while Section 4 presents the multivariate ones. Section 5 presents and discusses the results of applying the Perron and Yabu (2009b) structural change test and the Kim-Perron unit root test that allows for a structural break in the trend function. Also, the Bierens nonlinear co-trending test is applied to investigate the existence of a common nonlinear deterministic trend in temperatures and radiative forcing variables. Section 6 applies a rotated PCA to filter the common trend found by the co-trending test which allows to address the potential causes of the so-called hiatus, the recent slowdown in the rate of increase in temperatures. The existence of breaks in the slope of the warming trend is investigated both in observed radiative forcing and temperature series, as well as in the indices produced using the PCA. Section 7 summarizes the main findings.

2 Data.

The global (G), northern and southern hemispheric temperatures (NH, SH) data used in this paper come from the Climatic Research Unit's HadCRUT4 (G^H , NH^H , SH^H ; Morice et al., 2012) and the NASA database (G^N , NH^N , SH^N ; GISTEMP Team, 2015; Hansen et al., 2010). These databases differ basically in two aspects: 1) the way temperatures are extrapolated (or not) into regions without observing stations. This is particularly important since the HadCRUT4 excludes most of the Arctic, where the warming has been very large during the past decade and; 2) the datasets and methods used to adjust sea surface temperatures. Recently, it has been argued that some important characteristics of the observed warming such as the current slowdown in the warming could be artifacts due to how temperature data are processed (Karl et al., 2015). To investigate this possibility and to offer a sensitivity analysis for our results, the global temperature series from Berkeley

Earth (G^B) and from the dataset in Karl et al. (G^K ; 2015) are also considered. To represent the most important natural sources of inter-annual global and hemispheric climate variability we use the following indices (Enfield et al., 2001; Kerr, 2000; Hurrell, 1995; Wolter and Timlin, 1998): the Atlantic Multidecadal Oscillation (AMO); the Southern Oscillation Index (SOI), the North Atlantic Oscillation (NAO) and the Pacific Multidecadal Oscillation (PDO). These series are used to filter out the effects of natural variability oscillations on global and hemispheric temperature series. The radiative forcing series cover the period 1880-2011 and are available from the NASA Goddard Institute for Space Studies (Hansen et al., 2011). These series represent the effective radiative forcing which includes a number of rapid adjustments to the radiative imbalance (see Hansen et al., 2005). For the purposes of this paper, we use the well mixed greenhouse gases (WMGHG; carbon dioxide (CO₂), methane (CH₄), nitrous oxide (N₂O) and chlorofluorocarbons (CFCs)); the total radiative forcing (TRF) which includes WMGHG plus ozone (O₃), stratospheric water vapor (H₂O), solar irradiance, land use change, snow albedo, black carbon, reflective tropospheric aerosols and the indirect effect of aerosols; and the radiative forcing from stratospheric aerosols (STRAT). The temperature and radiative forcing series are presented in Figure 1. We adopt 1880-2014 as the common sample period ¹.

3 Tests and procedures for the univariate analyses.

The two nonstationary processes that have been mainly proposed for global and hemispheric temperatures are trend stationary (TS) and difference stationary (DS). The work of Tol and de Vos (1993; 1998) and Stern and Kauffman (2000) proposed the use of cointegration techniques to represent the long-term relationship between temperature and forcing variables. Nevertheless, results and inferences based on this technique depend on the presence of unit roots in temperature and forcing series, a proposition that could be only tested after recent advances in econometric modelling (see Perron, 2006 and Estrada and Perron, 2014) permeated the climate change literature (Gay et al., 2009; Estrada et al., 2013a,b).

The rate of warming observed during the 20th century can hardly be considered constant. The existence of breaks in global and hemispheric temperatures has been discussed extensively in the

¹All data are available at the following links: <http://www.metoffice.gov.uk/hadobs/hadcrut4/>;
<http://data.giss.nasa.gov/gistemp>; <http://berkeleyearth.org/land-and-ocean-data/>;
<ftp://ftp.ncdc.noaa.gov/pub/data/scpub201506/>; <http://www.esrl.noaa.gov/psd/data/timeseries/AMO/>;
http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/soi.long.data;
http://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/nao.long.data;
<https://www.ncdc.noaa.gov/teleconnections/pdo/data.csv>; http://data.giss.nasa.gov/modelforce/Fe_H11_1880-2011.txt

climate literature but it has been rarely formally investigated using tests that are adequate for the time-series properties of the series being analyzed (e.g., Seidel and Lanzante, 2004; Gay et al., 2009; Estrada et al., 2013b; IPCC, 2013). In the case of radiative forcing variables, even though it is commonly accepted that the atmospheric concentrations of greenhouse gases show different stages of growth during the past century, this has been only recently analyzed by means of formal structural break tests (e.g., Estrada et al., 2013b). If a linear trend provides an inadequate representation to describe the secular movement of temperature and forcing series, misleading results are to be expected both when analyzing their time-series properties (e.g. by means of unit root tests; Perron, 1989) as well as when conducting multivariate analyses (e.g., cointegration tests; Gonzalo and Lee, 1998). Consequently, a logical starting point to investigate the temperature and radiative forcing properties and the existence of shared secular trends (and other common features) is to expand the deterministic linear trend plus stationary noise to more realistic specifications of the trend function that allow the presence of nonlinearities. Below we describe the univariate tests used in this paper to investigate the univariate time-series properties of these series.

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

Perron (1989) showed that the presence of structural changes in the trend 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 and Yabu (2009b) 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, $I(0)$, or contains an autoregressive unit root, $I(1)$. The approach of Perron-Yabu builds on Perron and Yabu (2009a) who analyzed the problem of hypothesis testing on the slope coefficient of a linear trend when no information about the nature, $I(0)$ or $I(1)$, of the noise component is available.

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 Perron and Yabu (2009b). 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 Ψ is a $(r \times 1)$ vector of unknown parameters which are model specific and described below. The initial condition u_0 is assumed to be bounded in probability. The autoregressive coefficient is such that $-1 < \alpha \leq 1$ and therefore, both integrated and stationary errors are allowed. The interest is in testing the null hypothesis $R\Psi = \gamma$ where R is a $(q \times r)$ full rank matrix and γ 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 $T_B = [\lambda T]$ 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 $DT_t = (t - T_B)$ if $t > T_B$ and 0 otherwise 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 α when $\alpha = 1$. The estimate of α 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, see Andrews and Ploberger (1994). 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 α as suggested by Roy and Fuller (2001). 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 and Fuller (2001) biased corrected estimate $\hat{\alpha}_M$ (see Perron and Yabu (2009b) for the recommended specifications); 4) Apply the truncation $\hat{\alpha}_{MS} = \hat{\alpha}_M$ if $|\hat{\alpha}_M - 1| > T^{-1/2}$ and 1 otherwise; 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 $W_{FMS}(\lambda)$ to test for a break at date $T_B = [\lambda T]$; 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 $Exp-W_{FS} = \log [T^{-1} \sum_{\Lambda} \exp(W_{FMS}(\lambda)/2)]$ where

$\Lambda = \{\lambda; \epsilon \leq \lambda \leq 1 - \epsilon\}$ for some $\epsilon > 0$. We set $\epsilon = 0.15$ as is common the literature.

Note that the Perron-Yabu test can be performed sequentially, testing for the first break using the full sample and then testing for a second break in the subsamples defined by the estimated break date, and so on until a non-rejection.. This is justified from the results in Kejriwal and Perron (2010), which states that at each step of the iteration one can use the same critical value that applies to a one-break test.

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

As shown in Perron (1989), the estimate of the sum of the autoregressive coefficients in an autoregression of order p , say, is highly biased towards unity if there is a shift in the trend function unaccounted for. In this case, the unit root null is hardly rejected even if the series is composed of *i.i.d.* disturbances around the 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 (1989) proposed an extension of the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979, Said and Dickey, 1984) 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 (e.g., Box and Tiao, 1975), separating what can and cannot be explained by the noise in a time series. Our interest centers on 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:

$$y_t = \mu + \beta_1 t + \beta_2 DT_t + \tilde{y}_t \tag{2}$$

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

$$\tilde{y}_t = \alpha \tilde{y}_{t-1} + \sum_{i=1}^k c_i \Delta \tilde{y}_{t-i} + e_t \quad (3)$$

where the k lagged values of $\Delta \tilde{y}_{t-i}$ are added as a semi-parametric correction for autocorrelation. In the Perron (1989) test the break is assumed to occur at a known date. Later, Perron (1997) 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 by estimating the break date by minimizing the sum of squared residuals from regression (2). The resulting unit root test is then the t-statistic for testing that $\alpha = 1$ in regression (3) estimated by OLS. The critical values of the limit distribution of the test are tabulated in Perron (1997). See also, Zivot and Andrews (1992).

A problem with most procedures to test for a unit root 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. As a 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 and Perron (2009). Their procedure is based on a pre-test for a change in the trend function, namely the Perron and Yabu (2009b) test. If this pre-test rejects, the limit distribution of their modified unit root test is then the same as if the break date was known (Perron and Vogelsang, 1993). 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 by minimizing the sum of squared residuals using regression (2). Then construct a window around that estimate defined by a lower bound T_l and an upper bound T_h . A window of 10 observations was used. Note that, as shown by Kim and Perron (2009), 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 $t_\alpha(\hat{\lambda}_{tr}^{AO})$:

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

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_t^n .

3.3 Perron-Zhu methodology for constructing a confidence interval for the break date

Perron and Zhu (2005) 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. 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 (2005) 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. As has been proposed previously in the literature (e.g., Gay et al., 2009; Estrada et al., 2013a,b), the first specification with a stationary noise component is the most relevant to characterize the temperature and forcing variables discussed in this paper. The interested reader is referred to Perron and Zhu (2005) for the specifications and limiting distributions for the other models. The deterministic part is specified as:

$$d_t = \mu_1 + \beta_1 t + \beta_b DT_t.$$

Note that at the time of the break, the slope coefficient changes from β_1 to $\beta_1 + \beta_b$ but that the trend function is continuous at T_B . 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 $\{1, t, DT_t\}$, i.e., applying OLS to the model

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

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

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

where β_b^0 is the true value of the change in the slope parameter and σ^2 is the long-run variance of u_t estimated using the Bartlett kernel with Andrews' (1991) automatic bandwidth selection method using an AR(1) approximation. Note that the limiting distribution of the estimate of the break date does not depend on the structure of the errors (except from the long-run variance term σ^2). The limiting distribution does depend on the location of the break, having smaller variance as the break occurs closer to the middle of the sample. As expected, the variance decreases as the shift in the slope increases.

4 Tests and procedures for multivariate analyses.

In this section, we discuss the multivariate procedures used, namely Bierens' nonparametric nonlinear co-trending test and a rotated principal component analysis to separate common trends and natural variability.

4.1 Bierens nonparametric nonlinear co-trending test

Nonlinear co-trending is a special case of the more general "common features" concept described by Engle and Kozicki (1993). The advantage of the test proposed by Bierens (2000) is that the nonlinear trend does not have to be parameterized. The nonlinear trend stationarity model considered 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 a deterministic k -variate general nonlinear trend function that allow, in particular, structural changes. Nonlinear co-trending occurs when there exists a non-zero vector θ such that $\theta' f(t) = 0$. Hence, the null hypothesis of this test is that the multivariate time series z_t is nonlinear co-trending, 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, though the critical values would be different.

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

$$M_1 = T^{-1} \sum_{t=1}^T \hat{F}(t/T) \hat{F}((t/T))'$$

where $\hat{F}(x) = T^{-1} \sum_{t=1}^{[Tx]} (z_t - \hat{\beta}_0 - \hat{\beta}_1 t)$ if $x \in [T^{-1}, 1]$, $\hat{F}(x) = 0$ if $x \in [0, T^{-1})$ 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; also

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

where $m = T^\alpha$ with T the number of observations and $\alpha = 0.5$ as suggested by Bierens (2000).

Solving $|\hat{M}_1 - \lambda \hat{M}_2| = 0$ and denoting the r^{th} largest eigenvalue by $\hat{\lambda}_r$, the test statistic is $T^{1-\alpha} \hat{\lambda}_r$. 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 (2000). The existence of 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 indicates a strong secular co-movement in the $r + 1$ series.

4.2 Rotated principal component analysis (PCA) to separate common trends and natural variability.

PCA is commonly used to extract the main variability modes of a set of n interrelated variables and also to reduce dimensionality while retaining most of the variability present in the dataset (Jolliffe, 2002). This technique has been applied to both stationary and trending data for a wide range of purposes and has been extended to be applied to data with special structures. PCA has been widely used in climate sciences to separate the main modes of variability of climate data, to filter out noise and to obtain clearer signals and spatial patterns (e.g., Wilks, 2011; O'Lenic and Livezey, 1988; von Storch and Zwiers, 1999; von Storch and Navarra, 1999). It has also been extended by atmospheric scientists in various directions, notably for spatio-temporal data with evolving spatial patterns and cyclic statistics (e.g., Jolliffe, 2002; Kim and Wu, 1999; von Storch and Zwiers, 1999). PCA has been also commonly use in climate reconstructions and paleoclimate (e.g., Mann et al., 1998; Luterbacher et al., 2002; Evans et al., 2002). In econometrics, this technique has been used to develop methods for estimating and testing for common trends, both in the cointegration and cotrending context (e.g., Bai, 2004; Maddala and Kim, 1998; Hatanaka and Yamada, 2003; Harris, 1997; Stock and Watson, 1988, 2002).

The principal components Y_1, Y_2, \dots, Y_n are orthogonal linear combinations of the original dataset X of the form $Y_i = \sum_{j=1}^n a_{ij} x_j$. The first principal component is the linear combination $Y_1 = \sum_{j=1}^n a_{1j} x_j$ that maximizes $var(a_1' X) = a_1' \Sigma a_1$ subject to the constraint of $a_1' a_1 = 1$, where Σ is the variance-covariance matrix of X . According to results on maximizing quadratic forms on the unit sphere this is attained when a_1 is equal to the first eigenvector (i.e., the eigenvector that corresponds to the largest eigenvalue) of the variance-covariance matrix of X . The remaining

principal components are those linear combinations of $a'_j X$ that maximize $var(a'_j X)$ subject to the constraint $a'_j a_j = 1$ and $cov(a'_j X, a'_k X) = 0$ for all $j \neq k$. The PC or factor loadings L are calculated by multiplying the entries of the eigenvectors by the squared root of the eigenvalue that corresponds to that particular eigenvector and represent the correlation coefficients between the PCs and the original variables. The loadings are useful to determine which of the original variables are more closely associated to a particular PC.

To simplify the interpretation of the principal components and to further separate the variability modes in a set of data, the axis of the principal components can be rotated. Several rotation methods have been proposed to accomplish this objective and can be grouped in two main types: orthogonal and oblique rotation. In the case of rotated PCA, the PCs or factor scores are given by $F = BZ$, where F is the matrix of factor scores, Z is the matrix of standardized values of the original data, and $B = L(L'L)^{-1}$ is the matrix of factor score coefficients (Harman, 1976). In this paper, we propose the use of rotated PCA (varimax rotation normalized) to extract the principal modes of variation of temperature and radiative forcing variables, in particular their common trend mode. As has been discussed in the literature, natural low-frequency oscillations such as the AMO can exaggerate or mask the observed warming and therefore recent studies have suggested that it is convenient to filter out their effects when analyzing the features of global and northern hemisphere trends (Wu et al., 2011; Knudsen et al., 2011; Swanson et al., 2009; Estrada et al., 2013a). It is important to note that PCA has been applied to nonstationary data to extract trends and variability modes in different contexts. While Estrada et al. (2013a) filter out the effects of AMO from global and northern hemisphere temperatures by means of an OLS regression of temperatures on AMO, here we propose the use of rotated principal components on a set of variables composed of radiative forcing, temperature series and the main modes of natural climate variability. The objective of the rotated PCA analysis we present is to find the linear combinations that represent the estimates of the common trend component and the physical variability modes that are known to affect the noise around that trend, and that are orthogonal to each other. As such, these PCs allow to reexpress the original variables - which are correlated to different degrees (i.e., share common information)- as variables that represent linearly independent modes of variability (that is, to separate the variability modes in the set of variables). We believe that using this multivariate procedure, which considers the most important variables affecting global and hemispheric temperatures, will permit to extract an estimate of the warming trend less distorted by the effects of physical modes of variability.

5 Full-sample features of the temperatures and forcing series.

We now present the empirical results from applying the procedures discussed in Sections 3 and 4 to the various temperature and forcing series discussed in Section 2. We begin in Section 5.1 with a univariate analysis of temperature and radiative forcing series and establish the presence of common non-linear deterministic trends in Section 5.2. An in-depth analysis of the possible causes of the hiatus is presented in Section 6, where we discuss common explanations that have been advanced in the literature and how our results offer a very different perspective, especially with respect to the role of the natural variability modes in temperatures.

5.1 Univariate analysis of temperature and radiative forcing series

The results obtained by applying standard unit root tests to global and hemispheric temperatures seem to provide strong evidence for unit root processes (results not shown here). However, visual inspection of Figure 1 suggests the presence of structural breaks in the slopes of the trends of all temperatures series. This feature has been frequently reported in the literature (Jones et al., 1986a,b,c; Hansen and Lebedeff, 1987; IPCC, 2013; Gay et al., 2009). Standard unit root tests can be seriously affected by the presence of such breaks and their results could be misleading (Perron 1989, 2006).

For all the following results, the changing growth model is assumed for global and hemispheric temperature series given that global warming would imply a change in the rate of increase of these variables without a concurrent level shift (Gay et al., 2009; Estrada et al., 2013a). To test for the existence of such breaks we applied the testing procedure of Perron and Yabu (2009b) which is valid whether the noise component is $I(1)$ or $I(0)$ and thus circumvents the problem of pretesting for a unit root that is usually needed to implement structural change tests. According to the results of the Perron-Yabu test (Table 1), the stability of the slope parameter of the trend function for all temperature series is rejected at the 1% significance level, providing strong arguments for the need of unit root tests that allow for a one-time break in the trend function to investigate their time series properties.

For G and NH temperatures the estimates of the break dates are similar for both datasets: in the case of G the break dates are 1971 and 1978 for NASA and HadCRUT4, respectively, and 1982 for NH regardless of the dataset considered. These break dates are broadly similar to previous estimates in the literature (e.g., Gay et al., 2009; Estrada et al., 2013a). Note that, with the exception of SH^N , the confidence intervals of the break dates obtained using the Perron-Zhu procedure (Table 1) in general overlap with one another. As has been reported previously (e.g., Gay et al., 2009;

Estrada et al., 2013a), SH typically shows the largest differences in break date estimates when comparing different datasets: while for HadCRUT4 the break in the slope of G occurs in 1976, in the case of NASA the break occurs in 1925. Previous versions of NASA’s temperature dataset revealed the existence of a significant second break occurring in 1959. However, for the current version of NASA’s SH, although the data suggests the existence of another break in the slope in 1959, it is no longer statistically significant at any conventional levels. It is important to consider that the current version of NASA’s dataset uses the ERSST v4 sea surface reconstructions which includes significant changes in comparison to the ERSST v3 version (Huang et al., 2015; Liu et al., 2015). The low agreement between NASA and HadCRUT4 time series in the case of SH is remarkable. This can be illustrated by comparing the differences between NH and SH across datasets: while running an OLS regression of NH from NASA on NH from HadCRUT4 results in a R^2 value of 0.95, doing the same for SH leads to a R^2 value of just 0.78.

The total warming from 1880 to 2014 is estimated to be about 30% larger with HadCRUT4 than for NASA: 1.05°C and 1.44°C for G; 1.26°C and 1.64°C for NH; 0.88°C and 1.12°C for SH. The magnitudes of the pre- and post-break slopes reveal important differences from one dataset to the other, particularly in the case of the southern hemisphere for the pre-break period. SH from NASA is the only time series that shows a significant 46-year cooling trend (1880-1925). G and NH from both datasets and SH from HadCRUT4 show a moderate and slow pre-break warming of about 0.5°C per century. All temperature series experience an abrupt and large increase in the rate of warming. For NH the increase in the rate of warming is approximately five-fold, four-fold in the case of G and two-fold for SH from HadCRUT4. The post-break warming rate is considerably higher for HadCRUT4 than for NASA. These rates (per century) are 2.48°C and 3.37°C in the case of NH; 1.69°C and 2.55°C for G and; 0.99°C and 1.57°C for SH, respectively.

These large differences in some of the characteristics of the warming trend in G, NH and SH have been previously discussed in the literature (Gay et al., 2009; Swanson, 2009; Estrada et al., 2013a). According to these studies, the large contrast in warming and in break dates between hemispheres suggest that they may not be only the product of the climate system’s response to the increase in greenhouse gases (GHG) concentrations alone but also of dominant modes of internal variability that may mask the underlying warming trend. The next section shows that the differences in the trend functions reported here are the product of low-frequency oscillations; it also explores how different the underlying warming trends actually are and if they can be considered to be the same.

We apply the Kim-Perron test to investigate the type of data generating processes of the temperature variables. As shown in Table 1, the results provide clear evidence for trend-stationarity once a break in the slope of the trend function is allowed. For all temperature series, the null

hypothesis of a unit root is rejected at the 5% significance level once the breaks reported in Table 1 are taken into account. Thus, the unit root hypothesis is strongly rejected and the use of cointegration techniques is not appropriate to investigate the existence of common secular trends. These results suggest that there have been only a limited number of events that have altered the long-run path of global and hemispheric temperatures. These changes have manifested themselves at different times in global and hemispheric temperatures and have consisted in abrupt increases of different magnitudes in the rate of warming (Gay et al., 2009). Note that Table 1 also includes the results for the NASA southern hemisphere temperature series when choosing 1959 as the break date. The rejection of the null of a unit root is robust to this change in the break date.

While there has been a long debate about the time-series properties of global and hemispheric temperatures, radiative forcing variables have received less attention in this respect and have usually been assumed to be integrated processes when conducting attribution studies (Stern and Kaufmann, 2000; Tol and de Vos, 1993, 1998). Here, we analyze the time series properties of radiative forcing series using the same econometric techniques applied before to temperature series. According to the Perron-Yabu test both WMGHG and TRF show a highly significant break in the slope of their trend function occurring in 1960. The confidence intervals of these break dates are very narrow and do not overlap with those of the temperature series, probably due to the existence of natural variability oscillations (see the discussion in the following section). Both WMGHG and TRF show that the Earth's energy imbalance has been increasing since the last century and that the rate of this increase accelerated notably since 1960 as a result of the unprecedented and continuous global economic growth after World War II (see Estrada et al., 2013a). The post-break rate of growth in radiative forcing is about 4 times larger compared to the pre-break estimate in these series. Furthermore, the Kim-Perron tests strongly rejects the hypothesis of a unit root in both WMGHG and TRF and shows that both temperature and radiative forcing series share similar time-series properties. This enables the use of appropriate time-series models and tests to investigate the existence of common nonlinear deterministic secular trends. The commonly used assumption of forcing variables being integrated processes is not supported by the data. As shown in Estrada et al. (2103a; Supplementary Material S6), a system with broken-trend stationary variables that are co-trending is consistent with a zero dimensional energy balance model. It is also encompassing in the sense that it can explain the findings obtained when neglecting the breaks, namely that temperatures are deemed to be $I(1)$ and the forcings, such as the WMGHG, deemed to be $I(2)$ and not $I(1)$ as claimed by many studies such as Kaufmann et al. (2011); see Estrada et al. (2103a; Supplementary Material S8).

5.2 Common nonlinear deterministic trends

The univariate results obtained in the last section are broadly similar to those reported in the literature although significant changes have occurred in how temperature series are constructed (e.g., the new version of the sea surface temperatures dataset used by NASA). One of the most relevant aspect to test for is whether the conclusions regarding the existence of a common long-term trend between temperature and radiative forcing series are robust to the updates that have taken place in both temperature and radiative forcing datasets. To test for the existence of such a common nonlinear deterministic trend, we use the test of Bierens (2000) described in section 4.1. Tables 2a and 2b show the results of applying the Bierens test to two sets of time series: a) NASA’s G, NH, SH and WMGHG, TRF; b) HadCRUT4’s G, NH, SH and WMGHG, TRF. In both cases, the results indicate the existence of four co-trending vectors ($r = 4$) involving G, NH, SH, WMGHG and TRF and therefore strongly suggest the existence of one common nonlinear deterministic trend. Furthermore, as argued in Estrada et al. (2013a) these results suggest a dominant anthropogenic contribution to the observed warming: applying the nonlinear co-trending test to just WMGHG and TRF indicates the existence of a common trend in both radiative forcing variables; by construction this common trend can only come from WMGHG. Therefore, WMGHG must contain the common nonlinear trend in the two sets of temperature and forcing series described above. Although causality can hardly be established based only on statistical tests, attribution to anthropogenic forcing is strongly suggested when the results from the cotrending tests are combined with basic climate physics: WMGHG imparts the common nonlinear trend to TRF and in turn this common nonlinear trend is imparted to the global and hemispheric temperatures.

6 An analysis of the causes of the hiatus and the dating of the breaks

The existence and causes of the current slowdown in the warming —and the inability of current physical models to reproduce it— is one of the most active topics in climate change research (Tollefson, 2014, 2016). The range of hypotheses that have been proposed to explain the behavior of the warming trend during the recent decades can be summarized in three broad groups: the effects of natural variability; artefacts produced by temperature data biases and; changes in (or omission of) some of the natural and anthropogenic forcings. We provide further evidence on the existence of the recent slowdown in the warming and we confront some of the reasons in the literature that have been put forward to explain or deny the existence of the slowdown.

6.1 Competing hypotheses to explain the slowdown

The leading explanations involving natural variability are related to coupled ocean-atmosphere processes and heat exchange between oceans and atmosphere. According to Steinman et al. (2015), the combined effects of the low-frequency oscillations produced by AMO and PDO over the northern hemisphere temperatures were able to mask the warming trend for the past decade. The 1997-98 El Niño could have led the equatorial Pacific into a prolonged La Niña-like cold state by initially transferring enormous quantities of heat from the oceans to the atmosphere and subsequently cooled the Pacific Ocean (Kosaka and Xie, 2013). This prolonged cold state could have significantly contributed to a lower warming rate in global temperature. Trenberth and Fasullo (2013) proposed that about 30% of the heat produced by the Earth’s energy imbalance has been taken by the deep oceans due to changes in the Pacific surface winds and to a change to the negative phase of the PDO. Furthermore, it has been shown that inserting PDO-type patterns to global climate models can produce decade-long pauses in warming (Meehl et al., 2011).

The slowdown in the warming has also been explained as an artefact produced by biases in temperature data as well as by the lack of coverage at the poles and in Africa (Karl et al., 2015; Cowtan and Way, 2014; Curry, 2014). However, Fyfe et al. (2016) compared the trends in the bias-corrected data and those in the climate simulations included in the Climate Models Intercomparison Project (CMIP5) and found that the observed rate of warming since the beginning of the present century has been considerably lower than the average simulated rate. These authors concluded that the slowdown in the warming is a real phenomenon that cannot be discarded as a bias problem in the data.

Lower solar activity has also been identified as a possible contributor to the smaller rate of warming during the past decade (Lean and Rind, 2009), but it has been argued that this factor is unlikely to explain a substantial part of the reduced warming (Feulner and Rahmstorf, 2010). Increases in the atmospheric concentrations of aerosols due to the growth of coal consumption in Asia produced a negative radiative forcing that could partially explain the slowdown in the warming (Kaufmann et al., 2011). Another factor that could explain the differences in the observed and simulated rates of warming is the volcanic aerosol forcing which is not well represented in CMIP5 (Andersson et al., 2015). Estrada et al. (2013a) suggest that part of the slowdown in the rate of warming can be directly linked to the effects of the Montreal Protocol and changes in agricultural production in Asia on the atmospheric concentrations of CFCs (Velders et al., 2007) and methane (Kai et al., 2011).

Table 3 updates some of the results in Estrada et al. (2013a) and confirms the existence of a slowdown in WMGHG and TRF. The methods used follows the Perron-Yabu sequential procedure

as proposed by Kejriwal and Perron (2010). As discussed above, both WMGHG and TRF are characterized by a highly significant break in 1960 (Table 1), and also by a second break occurring in 1994 and 1991, respectively. In both cases the second break in the slope of their trend function is significant at the 1% level (Table 3). This finding provides strong evidence for the existence of a slowdown in the underlying common trend and for its anthropogenic origins². The rate of growth of WMGHG after the break decreased about 25%, while that of TRF decreased about 56% due to the decrease in WMGHG and the effects of changes in other forcing factors such as atmospheric aerosols.

We now turn to various observed global temperature series to test for a break in the slope of their trend function that could be consistent with a slowdown in the warming during the last two decades. In addition to the NASA and HadCRUT4 global temperature series, this subsection considers also those of Berkeley Earth and of the dataset in Karl et al. (2015). These two additional datasets are of interest to our study for the following reasons: Karl et al. (2015) claimed that the pause in the warming is an artefact of data biases, and they produced a series labelled G^K which is arguably free of such biases; according to the Berkeley Earth project they offer a more complete sampling using five times more data than other groups and follow an independent approach based on modern statistical techniques. As shown in Table 3, both G^H and G^B (two out of four of these global temperature series) do show a significant break in their slope around the mid-2000 that is consistent with the slowdown in the warming that has been widely discussed in the literature (the decrease in the warming rate after the break dates is about 50%). For G^N and G^K no evidence of a second break can be found. By the time of writing this paper, the G^H and G^B observed temperature values for 2015 became available. Including this extra observation makes the second break in both G^H and G^B no longer significant. This finding needs to be interpreted carefully as 2015 was an extraordinarily warm year due in part to one of the strongest El Niño events on record. As such, the 2015 record temperature value may not contribute much to clarifying the features of the underlying warming trend. A tempting conclusion would be that the slowdown never occurred. However, this conclusion would be difficult to support given the evidence about the slowdown in the radiative forcing series discussed above and in Estrada et al. (2013a).

²A sensitivity analysis was conducted using the dataset of Mills et al. (2014) and the results are robust to this alternative radiative forcing estimates. For both TRF and WMGHG, highly significant breaks were found to occur in the trend function. The first break date occurs around 1960 (1962 for TRF and 1959 for WMGHG), while other statistically significant breaks were found in the post-1960 sample in 1990 for TRF and 1991 for WMGHG.

6.2 Filtering procedure based on a rotated principal component analysis

Observation-based attribution studies have been mainly limited to establishing the existence of a common long-term trend between temperature and radiative forcing variables. Recent work have proposed to advance the study of the attribution of climate change by focusing on characterizing the common secular trend (Estrada et al., 2013a; Estrada and Perron, 2014). This allows investigating the reasons behind periods of fast warming, slowdowns and pauses but it requires the extraction of the underlying secular trend.

Estrada et al. (2013a) analyzed the characteristics of the common secular trend in temperature and radiative forcing by: 1) filtering out the low-frequency oscillations produced by AMO that could distort the warming trend in G and NH temperatures by means of a simple OLS regression and; 2) investigating the existence of breaks in the common trend in temperatures and radiative forcing series directly in WMGHG and TRF. The main advantage of this approach is that radiative forcing series are considerably less noisy than temperature data and are known to be closely related to the transient climate response. Global surface temperatures can be expressed as $T_t = \alpha + \gamma F_t + \varepsilon_t$ where T_t is temperature, F_t is a measure of the radiative forcing, ε_t encompasses both short- and long-term natural variability and γ represents the transient climate sensitivity (see Estrada et al., 2013a; Schwartz, 2012; Gregory and Forrester, 2008). The structural model behind the time series models in our paper can be described by a two-compartment model of the climate system (see Schwartz, 2012; Estrada et al., 2013a). The upper compartment includes the atmosphere and the upper ocean and it is characterized by a small heat capacity and a short time constant (of about a few years) for reaching steady state following a perturbation. In contrast, the lower compartment—which represents the deep ocean—has a large heat capacity and a long time constant (in the order of hundreds of years) for reaching a steady state. The response of the climate system to changes in external forcing during the observed period is determined by the time constant of the upper compartment and the transient climate sensitivity. This simple structural model helps understanding from a physical perspective why global and hemispheric surface temperatures follow the same nonlinear trend of the radiative forcing and also the rapid adjustment of observed temperatures to changes in the trend of the radiative forcing (see Estrada et al., 2013a, Supplementary Information S6).

Based on the finding of a common nonlinear long-term trend in temperature and radiative forcing variables and on the literature on climate variability, we propose the use of a rotated principal component analysis to separate the warming trend from other modes of variability present in the data. In this way, the features of the common secular trend such as the sharp increase in the warming since the mid-20th century and the existence of the slowdown in the warming of the last decades can be better analyzed. The set of variables selected for the rotated principal component

analysis include: G, NH, SH, WMGHG, TRF which according to the results of the co-trending analysis contain a common nonlinear trend; PDO (Zhang et al., 1997), AMO (Enfield et al., 2001), SOI (Trenberth, 1984) and NAO (Hurrell, 1995) which are commonly considered as some of the most important natural sources of inter-annual global and hemispheric climate variability (IPCC, 2013; Wolter and Timlin, 1998; Estrada et al., 2013a) and the radiative forcing from stratospheric aerosols produced by volcanic eruptions (STRAT). Tables 4a and 4b show the factor loadings for the sets containing G, SH and NH from the NASA and HadCRUT4 datasets, respectively.

The results presented below are based on extracting and rotating the ten possible principal components. It is important to note that the application of PCA in this paper is not to reduce dimensionality but to separate the modes of variability and, in particular, to obtain an estimate of the climate trend that is less noisy than temperatures and free of the major sources of natural variations. Different objectives for a PCA lead to different recommendations concerning how many PCs to retain and, in some cases, the size of the eigenvalues has no relationship with which PCs are of interest (Jolliffe, 2002); these could be very well beyond what some truncation rule could suggest (see Preisendorfer and Mobley, 1982, 1988). Furthermore, for some applications such as the regression analysis presented in this section, it is inadvisable to look only at high-variance PCs, as the low-variance PCs can also be correlated with and thereby explain the dependent variable (Jolliffe, 2002). As shown below, some of the low-variance PCs contribute significantly to explain G, NH and SH and the differences in break dates. When applying the rotated PCA, retaining too few or too many PCs, underrotation or overrotation, can be problematic. In the first case, part of the signal is unjustifiably discarded, while in the second case unnecessary noise can be introduced. In general, the effects of underrotation are considered more serious than those of overrotation, given that part of the signal is being discarded and this is often accompanied by the distortion of patterns (e.g., O’Lenic and Livezey, 1988). The screen plot of the PCA suggests that the truncation point should be after the 5th or 6th eigenvector. As suggested by O’Lenic and Livezey (1988) and Richman (1981) it is preferable to retain slightly more modes in order to avoid underrotation, so the rotation point was chosen to be the 6th PC. The results presented in this section are robust to this truncation point: the eigenvectors entries from the untruncated/truncated rotated PCA are essentially the same and the corresponding PCs are indistinguishable from one another (the correlation coefficients are in all cases above 0.99).

The first principal component (PC1) is highly correlated with G, NH, SH, WMGHG, TRF (all correlation coefficients are equal to or larger than 0.90) and has almost zero correlation with all other variables. Figure 2 shows that PC1 is the trend mode of the series in datasets. It should be noted that PC1 is intended to be viewed as an estimate of the underlying trend and, hence,

as for all estimates, it is subject to errors (and, here also somewhat influenced by the number of components used). What is important for our purpose is that it is trending, shows less noisy than temperature series (which it is) and be free of the main modes of natural variability.

Each of the next 5 modes are highly correlated (>0.95) with one and only one of the natural variability series included in the analysis: PC2 with PDO; PC3 with STRAT; PC4 with NAO; PC5 with AMO; PC6 with SOI. Tables 5a and 5b show that these first six principal components account for about 98% of the total variance of the dataset (the rotated first principal component itself accounts for about 46% of the variability). PC8 from NASA and PC7 from HadCRUT4 are grouped together in Figure 2 since they represent the same low-frequency variability mode, likely related to other modes such as the Interdecadal Pacific Oscillation (IPO) and the Southern Annular Mode (SAM). As discussed below, PC7 and PC8 obtained from NASA and HadCRUT4, respectively, are mainly the particular variability mode that make SH from NASA and HadCRUT4 so different from each other. PC9 in both datasets recovers the solar cycle originally included in TRF. With the exception of PC1, according to standard unit root tests (not shown here) all principal components can be considered stationary processes around a constant which further confirms that PC1 represents the only trending component. Previous studies have applied PCA to separate the main modes of variability in global temperature datasets (e.g., Mann et al., 1998). In such studies, PCA was applied to surface temperature datasets that have a spatio-temporal structure. Therefore, in such cases the eigenvectors refer to the spatial variability of the analyzed variables. Here, we are dealing with a set of different but interrelated variables in which there is only a temporal dimension. The eigenvectors in this case refer to temporal variability. The results presented here are broadly similar to those of previous studies in terms of the relative importance of the variability modes.

Results about the statistical evidence of a change in slope based on the Perron-Yabu test, henceforth PY, and the estimate of the break date, are presented in Table 3 for PC1 and all combinations of PC1 with another principal component. PC1 is characterized by highly significant breaks in its slope during the 1960s and the 1990s (the first break occurs in 1962 and 1968, and the second break in 1989 and 1991 for NASA and HadCRUT4, respectively). These break dates are not statistically different from those of TRF. The increase in the rate of warming decreased after the second break dates about 31% and 40% in the cases of PC1 from HadCRUT4 and NASA, respectively. As expected from climate physics and from the empirical evidence on co-trending, the recovered warming trend is similar to the nonlinear trend in forcing once the natural variability noise component is adequately removed. These results suggest that the proposed method is capable of adequately separating the underlying nonlinear trend function present in both temperature and

forcing variables from modes of natural variability. The recovered modes are also useful to better understand the main variability factors in G, NH and SH. For example, PC1 and PC5 (AMO) are the two most important factors dominating G and NH and explain about 95% and 92% of their variability, respectively, for both NASA and HadCRUT4 datasets³. The variability mode represented by PC7 and PC8 in the results for NASA and HadCRUT4, respectively, and the warming trend (PC1) are able to explain about 95% of the variability of SH.

As discussed in what follows, these variability modes (PC5, PC7 (NASA) and PC8 (HadCRUT4)) are to a large extent responsible for the differences in the break date estimates in G, NH and SH and between datasets, and they are also responsible for masking features of the underlying common trend (e.g., the current slowdown in the warming). To investigate this, regressions can be estimated using G, NH and SH as dependent variables and different combinations of PCs as independent variables. These regressions are used to estimate the coefficients to scale the PC scores to reconstruct the original series (e.g., G, NH, SH). By construction, if all PCs were included in the regression the explained variance would be 100%, and the series would be exactly reproduced. Given that PCs are orthogonal the estimated coefficients are unaffected if other PCs are excluded (there is no “omitted variables” problem). Using the sum of PC1 and PC5, scaled by the estimated regression coefficients, changes the date of the first break to 1974 (NASA) and 1975 (HadCRUT4), which are not statistically different from the ones obtained from the original NH series. Furthermore, once the effects of PC5 (AMO) are added to PC1, the slowdown in the warming is no longer detectable (PY values of 0.27 and 0.96, respectively). In the case of G, using the weighted sum of PC1 and PC5 according to the estimated regression coefficients changes the break date estimate to 1971 (NASA) and 1975 (HadCRUT4), which are not statistically different from those estimated for the original series (Table 1). The slowdown in the warming is not detectable anymore in either NASA or HadCRUT4 (PY values of 0.26 and 0.98, respectively). For SH from NASA, once the effects of the variability mode PC7 are added to the warming trend, the estimated break date is 1929, which is not statistically different from the break date in the original series, and the slowdown in the warming is not detectable (PY value of -0.13). In the case of SH from HadCRUT4, adding the effects of PC8 to the warming trend changes the estimated break date to 1967, a date that is not statistically different from the one in the original temperature series. As in the other cases, the addition of this variability mode makes the slowdown in the warming not detectable anymore (PY value of -0.07). It is important to notice that while for G and NH, PC5 constitutes a clear natural variability mode (AMO) which is similar for both datasets (the correlation coefficient between PC5

³The estimates of how much of the variance of a particular variable is explained by different combinations of PCs can be obtained by squaring the corresponding loadings given in Tables 4a and 4b and adding them. Alternatively, these estimates can be obtained from the R^2 obtained from regressing a particular variable on the PCs of interest.

from NASA and HadCRUT4 is 0.999), this is not the case for PC7 from NASA and PC8 from HadCRUT4. PC7 and PC8 do represent the same mode of variability in the different datasets, but given their dissimilarities (e.g., their correlation coefficient is 0.31), the question remains whether these series are related to a real natural variability mode or if they just reflect the differences in how SH series are constructed (e.g., differences in data coverage and interpolation, differences in processing and adjusting ocean data).

These findings show how natural variability can substantially change the features in the estimated underlying warming trend and makes it seem that G, NH and SH have responded very differently to the observed changes in the radiative forcing. These differences largely disappear when these modes are included in the analysis and the warming trend becomes very similar among G, NH and SH and consistent with the features found in TRF. That is, the results strongly suggest that all the differences in the break date estimates and other characteristics reported in the previous sections can be explained by natural variability modes and the differences in how temperature data are reconstructed. In the same way, testing for the existence of the recent slowdown in global temperatures without taking into account the effects of natural variability is likely not to provide much information about the evolution of the underlying warming trend. This is one of the main reasons that could explain why recent studies have not found statistical evidence for the existence of a slowdown in warming (Karl et al., 2015; Lewandowsky et al., 2016) and why probably many more will conclude the same after including the 2015 record value in their analyses. It is important to recognize that the question “is there a slowdown in the warming trend?” is different from “is there a slowdown in global temperature series?” Natural variability modes play a large role in defining the evolution of the observed global temperature series (e.g., Trenberth, 2015) and ignoring this fact can lead to erroneous conclusions regarding the warming trend and its features.

Our results show that natural variability modes such as PDO and AMO cannot account for the current slowdown in the warming trend as has been proposed frequently in the literature. The empirical evidence indicates that the slowdown comes from changes in the radiative forcing. The current slowdown is driven by anthropogenic factors and illustrates how dominant human activities can be in defining climate in different time-scales. In addition, the results also suggest that large methodological differences on how ocean temperatures are reconstructed and adjusted may be important factors making it difficult to find consistent results between SH and NH and the different datasets that are available.

It is important to put our results in perspective within the recent literature. As discussed in Section 6.1, the main contenders to explain the so-called hiatus are: natural variability; data biases; changes in anthropogenic forcing (mostly CFC and methane reductions). Our results provide

evidence against natural variability and data biases and in favor of changes in anthropogenic forcings. First, there is a clear reduction in the slope of PC1, associated with the estimated trend, induced by anthropogenic forcings. This is the argument advanced by Estrada et al. (2013a) who established that the reduction was mostly due to reduced CFC and methane radiative forcing, as well as increases in atmospheric aerosols. Of all the various other components, only PC5 and PC7 can mask the slowdown (as well as PC3 in the case of the HadCRUT4 series). For both the NASA and HadCRUT4 series PC5 is the AMO, while PC7 is a variability mode likely related to IPO or SAM for HadCRUT4, while for NASA PC7 has to do with adjustments regarding the southern hemisphere ocean temperatures (closely related to the arguments advanced in Karl et al., 2015). Hence, in direct contrast to what has been advanced in the literature, our results indicate that the changes in the anthropogenic forcings are directly responsible for the hiatus, while the following contribute to blurring the signal and making the statistical evidence weaker: the AMO (the main one), IPO or SAM, and different temperature adjustments such as those in Karl et al. (2015). In other words, natural variability and data adjustments do not explain in any way the hiatus, they simply mask its presence. We believe these results should spur additional debates about the causes of the hiatus. The results are also able to reconcile the differences in the estimates of the break dates across various series and across global and hemispheric temperatures and strongly suggest that these differences are caused by natural variability as well as differences in how temperature data are reconstructed and adjusted.

7 Conclusions

The attribution of climate change to anthropogenic activities presented in this paper takes into account natural and anthropogenic forcing variables as well as the most important natural variability modes referred to in the literature, and uses a variety of temperature records. State-of-the-art econometric techniques strongly suggest that both temperature and radiative forcing variables are better represented as trend stationary variables with breaks, that they share the same order of integration and, more importantly, that temperatures and forcing variables share the same long-term trend which is shown to be highly influenced by the well-mixed greenhouse gases forcing. The empirical evidence indicates that while the main modes of climate variability can impart low-frequency oscillations to global and hemispheric temperatures that can mask or exaggerate the warming, there recently was a slowdown in the underlying warming trend that is directly related to the slowdown in the total radiative forcing discussed in Estrada et al. (2013a). The results presented here also offer a simple explanation for at least part of the discrepancy between the observed rate of warming and the average rate of warming in the CMIP5 simulations. The radiative

forcing used in the CMIP5 experiments was constructed using estimates of observed quantities during the period 1765-2005 and projected quantities corresponding to the RCP4.5 scenario from 2006 onwards (Meinshausen et al., 2011), effectively eliminating the slowdown in the observed total radiative forcing. The results show that the slowdown is a characteristic of the underlying warming trend that is present at least until 2011. However, observed data on radiative forcing since 2011 is needed to assess whether the slowdown in the warming is still present.

The available literature on attribution has shown that the finding of a common long-term trend in temperatures and radiative forcing variables is robust to a wide variety of assumptions and methods (both statistical and physical). Although this literature is characterized by large and important methodological debates, the attribution of climate change to human activities seems to be a settled issue. This paper follows Estrada et al. (2013a) in advancing attribution studies towards the characterization of the underlying warming trend, which can provide relevant information on how the climate system has responded to changes in radiative forcing and also on how it could respond to international mitigation agreements. Using state-of-the-art econometric techniques, the warming trend is isolated and some of its predominant features are characterized. This paper contributes to a better understanding of the role of natural variability oscillations in masking the underlying warming trend and suggests that differences in how southern hemisphere temperatures are reconstructed could be distorting this trend. Furthermore, it contributes to the debate regarding the current slowdown in the warming and provides further evidence of its existence and origins. It is shown that due to the effects of climate variability over the warming trend, multivariate signal extraction techniques are useful to extend current attribution studies. To that effect, a rotated principal components analysis was applied. As is common in empirical analyses of observed data, the results may be dependent on the filtering method and the data sample. In particular, the radiative forcing series ends 2011 and having more recent data for this variable could help to better characterize the warming trend.

Our analysis offers some provocative results in relation to the recent literature. Changes in the anthropogenic forcings are directly responsible for the hiatus, while the AMO (especially), the IPO or SAM, and the new temperature adjustments in Karl et al. (2015) contribute to blurring the signal and making the statistical evidence weaker. In other words, natural variability and data adjustments do not explain in any way the hiatus, they simply mask its presence. As shown by the results of the PCA, natural variability and differences in temperature data reconstruction and adjustment procedures are also able to explain the difference in the estimates of the break dates across various series and across global and hemispheric temperatures.

References

- Andersson SM, Martinsson BG, Vernier JP, Friberg J, Brenninkmeijer CAM, Hermann M, van Velthoven PFJ, Zahn A. 2015. Significant radiative impact of volcanic aerosol in the lowermost stratosphere. *Nature Communications* **6**: 7692.
- Andrews DWK. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**: 817-858.
- Andrews DWK, Ploberger W. 1994. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* **62**: 1383-1414.
- Bai J. 2004. Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics* **122**: 137-183.
- Bierens HJ. 2000. Nonparametric nonlinear cotrending analysis, with an application to interest and inflation in the United States. *Journal of Business and Economic Statistics* **18**: 323-337.
- Box GEP, Tiao GC. 1975. Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association* **70**: 70-79.
- Cowtan K, Way RG. 2014. Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Quarterly Journal of the Royal Meteorological Society* **140**: 1935-1944.
- Curry J. 2014. Climate science: Uncertain temperature trend. *Nature Geoscience* **7**: 83-84.
- Dickey DA, Fuller WA. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* **74**: 427-431.
- Enfield DB, Mestas-Nunez AM, Trimble PJ. 2001. The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the continental U.S. *Geophysical Research Letters* **28**: 2077-2080.
- Engle RF, Granger CWJ. 1987. Co-integration and error correction: representation, estimation and testing. *Econometrica* **55**: 251-276.
- Engle RF, Kozicki S. 1993. Testing for common features. *Journal of Business and Economic Statistics* **11**: 369-395.
- Estrada F, Gay C, Sánchez A. 2010. Reply to ‘Does temperature contain a stochastic trend? Evaluating conflicting results by Kaufmann et al. *Climatic Change* **101**: 407-414.
- Estrada F, Perron P. 2014. Detection and attribution of climate change through econometric methods. *Boletín de la Sociedad Matemática Mexicana* **20**: 107-136.
- Estrada F, Perron P, Martínez-López B. 2013a. Statistically derived contributions of diverse human influences to twentieth-century temperature changes. *Nature Geoscience* **6**: 1050-1055.

- Estrada F, Perron P, Gay-García C, Martínez-López B. 2013b. A time-series analysis of the 20th century climate simulations produced for the IPCC's Fourth Assessment Report. *Plos One* **8**(3): e60017.
- Evans MN, Kaplan A, Cane MA. 2002. Pacific sea surface temperature field reconstruction from coral ^{18}O data using reduced space objective analysis. *Paleoceanography* **17**: 7(1)-7(13).
- Feulner G, Rahmstorf S. 2010. On the effect of a new grand minimum of solar activity on the future climate on Earth. *Geophysical Research Letters* **37**: L05707.
- Fyfe JC, Meehl GA, England MH, Mann ME, Santer BD, Flato GM, Hawkins E, Gillett NP, Xie SP, Kosaka Y, Swart NC. 2016. Making sense of the early-2000s warming slowdown. *Nature Climate Change* **6**: 224-228.
- Gay C, Estrada F, Sanchez A. 2009. Global and hemispheric temperature revisited. *Climatic Change* **94**: 333-349.
- Gil-Alana LA. 2008. Time trend estimation with breaks in temperature time series. *Climatic Change* **89**: 325-337.
- GISTEMP Team. 2015. GISS surface temperature analysis (GISTEMP). NASA Goddard Institute for Space Studies. Dataset accessed 2016-01-08 at <http://data.giss.nasa.gov/gistemp/>.
- Gonzalo J, Lee TH. 1998. Pitfalls in testing for long run relationships. *Journal of Econometrics* **86**: 129-154.
- Gregory JM, Forster PM. 2008. Transient climate response estimated from radiative forcing and observed temperature change. *Journal of Geophysical Research* **113**: D23105.
- Hansen J, Sato M, Kharecha P, von Schuckmann K. 2011. Earth's energy imbalance and implications. *Atmospheric Chemistry and Physics* **11**: 13421-13449.
- Hansen J, Ruedy R, Sato M, Lo K. 2010. Global surface temperature change. *Reviews of Geophysics* **48**: RG4004.
- Hansen J, Lebedeff S. 1987. Global trends of measured surface air temperature. *Journal of Geophysical Research* **92**: 13345-13372.
- Hansen J, Sato MK, Ruedy R, Nazarenko L, Lacis A, Schmidt GA, Russell G, Aleinov I, Bauer M, Bauer S, Bell N. 2005. Efficacy of climate forcings. *Journal of Geophysical Research: Atmospheres* **110**: D18104.
- Harman H. 1976. *Modern factor analysis*. University of Chicago Press, Chicago, USA.
- Harris D. 1997. Principal components analysis of cointegrated time series. *Econometric Theory* **13**: 529-557.
- Hasselmann K. 1979. On the signal-to-noise problem in atmospheric response studies. In: *Meteorology of Tropical Oceans*, Shaw DB (ed.): Royal Meteorological Society, Bracknell, UK, 251-259.

- Hasselmann K. 1997. Multi-pattern fingerprint method for detection and attribution of climate change. *Climate Dynamics* **13**: 601-612.
- Hatanaka M, Yamada H. 2003. *Co-trending: A statistical system analysis of economic trends*. Springer Science & Business Media, Tokyo, Japan.
- Huang B, Banzon VF, Freeman E, Lawrimore J, Liu W, Peterson TC, Smith TM, Thorne PW, Woodruff SD, Zhang H. 2015. Extended reconstructed sea surface temperature version 4 (ERSST.v4). Part I: Upgrades and intercomparisons. *Journal of Climate* **28**: 911-930.
- Hurrell JW. 1995. Decadal trends in the North Atlantic Oscillation and relationships to regional temperature and precipitation. *Science* **269**: 676-679.
- IPCC. 2013. *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds.): Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- IPCC. 2014. *Climate change 2014: Impacts, adaptation, and vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, White LL (eds.): Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Jolliffe IT. 2002. *Principal Component Analysis*. John Wiley & Sons, New York.
- Jones PD, Raper SCB, Bradley RS, Diaz HF, Kelly PM, Wigley TML. 1986a. Northern hemisphere surface air temperature variations: 1851-1984. *Journal of Climate and Applied Meteorology* **25**: 161-179.
- Jones PD, Raper SCB, Wigley TML. 1986b. Southern hemisphere surface air temperature variations: 1851-1984. *Journal of Climate and Applied Meteorology* **25**: 1213-1230.
- Jones PD, Wigley TML, Wright PB. 1986c. Global temperature variations between 1861 and 1984. *Nature* **322**: 430-434.
- Kai FU, Tyler SC, Randerson JT, Blake DR. 2011. Reduced methane growth rate explained by decreased northern hemisphere microbial sources. *Nature* **476**: 194-197.
- Karl TR, Arguez A, Huang B, Lawrimore JH, McMahon JR, Menne MJ, Peterson TC, Vose RS, Zhang HM. 2015. Possible artifacts of data biases in the recent global surface warming hiatus. *Science* **348**: 1469-1472.
- Kejriwal M, Perron P. 2010. A sequential procedure to determine the number of breaks in trend with an integrated or stationary noise component. *Journal of Time Series Analysis* **31**: 305-328.
- Kerr RA. 2000. A North Atlantic climate pacemaker for the centuries. *Science* **288**: 1984-1985.

- Kaufmann RK, Kauppi H, Mann ML, Stock JH. 2011. Reconciling anthropogenic climate change with observed temperature 1998-2008. *Proceedings of the National Academy of Sciences* **108**: 11790–11793.
- Kim D, Perron P. 2009. Unit root tests allowing for a break in the trend function under both the null and the alternative hypotheses. *Journal of Econometrics* **148**: 1-13.
- Kim KY, Wu Q. 1999. A comparison study of EOF techniques: Analysis of nonstationary data with periodic statistics. *Journal of Climate* **12**: 185-199.
- Knudsen MF, Seidenkrantz MS, Jacobsen BH, Kuijpers A. 2011. Tracking the Atlantic Multidecadal Oscillation through the last 8,000 years. *Nature Communications* **2**: 178.
- Kosaka Y, Xie SP. 2013. Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature* **501**: 403–407.
- Lean JL, Rind DH. 2009. How will Earth’s surface temperature change in future decades? *Geophysical Research Letters* **36**(15): L15708.
- Lewandowsky S, Risbey JS, Oreskes N. 2016. The “pause” in global warming: Turning a routine fluctuation into a problem for science. *Bulletin of the American Meteorological Society* **97**: 723–733.
- Liu W, Huang B, Thorne PW, Banzon VF, Zhang HM, Freeman E, Lawrimore J, Peterson TC, Smith TM, Woodruff SD. 2015. Extended reconstructed sea surface temperature version 4 (ERSST.v4). Part II: Parametric and structural uncertainty estimations. *Journal of Climate* **28**: 931–951.
- Luterbacher J, Xoplaki E, Dietrich D, Rickli R, Jacobeit J, Beck C, Gyalistras D, Schmutz C, Wanner H. 2002. Reconstruction of sea level pressure fields over the Eastern North Atlantic and Europe back to 1500. *Climate Dynamics* **18**: 545–561.
- Maddala GS, Kim IM. 1998. *Unit roots, Cointegration, and Structural Change*. Cambridge University Press, Cambridge, UK.
- Mann ME, Bradley RS, Hughes MK. 1998. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392**: 779-787.
- Meehl GA, Arblaster JM, Fasullo JT, Hu A, Trenberth KE. 2011. Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Climate Change* **1**: 360–364.
- Meinshausen M, Smith SJ, Calvin K, Daniel JS, Kainuma MLT, Lamarque JF, Matsumoto K, Montzka SA, Raper SCB, Riahi K, Thomson A, Velders GJM, van Vuuren DPP. 2011. The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change* **109**: 213–241.
- Miller RL, Schmidt GA, Nazarenko LS, Tausnev N, Bauer SE, DelGenio AD, Kelley M, Lo KK, Ruedy R, Shindell DT, Aleinov I. 2014. CMIP5 historical simulations (1850–2012) with GISS ModelE2. *Journal of Advances in Modeling Earth Systems* **6**: 441-477.

- Mitchell JFB, et al. 2001. Detection of climate change and attribution of causes. In: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, Houghton JT, et al. (eds.): Cambridge University Press, Cambridge, United Kingdom and New York, NY, 695-738.
- Morice CP, Kennedy JJ, Rayner NA, Jones PD. 2012. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 dataset. *Journal of Geophysical Research* **117**: D08101.
- O'Lenic EA, Livezey RE. 1988. Practical considerations in the use of rotated principal component analysis (RPCA) in diagnostic studies of upper-air height fields. *Monthly Weather Review* **116**: 1682-1689.
- Perron P. 1989. The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* **57**: 1361-1401.
- Perron P. 2006. Dealing with structural breaks. In *Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory*, Patterson K, Mills TC (eds.): Palgrave Macmillan, New York, 278-352.
- Perron P. 1997. Further evidence on breaking trend functions in macroeconomic variables. *Journal of Econometrics* **80**: 355-385.
- Perron P, Vogelsang TJ. 1993. Erratum: The great crash, the oil price shock and the unit root hypothesis. *Econometrica* **61**: 248-249.
- Perron P, Yabu T. 2009a. Estimating deterministic trends with an integrated or stationary noise component. *Journal of Econometrics* **151**: 56-69.
- Perron P, Yabu T. 2009b. Testing for shifts in trend with an integrated or stationary noise component. *Journal of Business and Economic Statistics* **27**: 369-396.
- Perron P, Zhu X. 2005. Structural breaks with deterministic and stochastic trends. *Journal of Econometrics* **129**: 65-119.
- Preisendorfer RW, Mobley CD. 1988. *Principal Component Analysis in Meteorology and Oceanography*. Elsevier Science, Amsterdam.
- Preisendorfer RW, Mobley CD. 1982. *Climate Forecast Verifications, U. S. Mainland, 1974-82*. NOAA Technical Memorandum ERL PMEL-36.
- Richman M., 1981. Obliquely rotated principal components: an improved meteorological map typing technique. *Journal of Applied Meteorology* **29**: 1145-1149.
- Rosenzweig C, Karoly D, Vicarelli M, Neofotis P, Wu Q, Casassa G, Menzel A, Root TL, Estrella N, Seguin B, Tryjanowski P, Liu C, Rawlins S, Imeson A. 2008. Attributing physical and biological impacts to anthropogenic climate change. *Nature* **453**: 353-357.
- Roy A, Fuller WA. 2001. Estimation for autoregressive processes with a root near one. *Journal of Business and Economic Statistics* **19**: 482-493.

- Said SE, Dickey DA. 1984. Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika* **71**: 599-607.
- Schwartz SE. 2012. Determination of Earth's transient and equilibrium climate sensitivities from observations over the twentieth century: strong dependence on assumed forcing. *Surveys in Geophysics* **33**: 745-777.
- Seidel DJ, Lanzante JR. 2004. An assessment of three alternatives to linear trends for characterizing global atmospheric temperature changes. *Journal of Geophysical Research* **109**: D14108.
- Steinman BA, Mann ME, Miller SK. 2015. Atlantic and Pacific multidecadal oscillations and Northern Hemisphere temperatures. *Science* **347**: 988-991.
- Stern DI, Kaufmann RK. 2000. Is there a global warming signal in hemispheric temperature series: a structural time series approach. *Climatic Change* **47**: 411-438.
- Stock JH, Watson MW. 1988. Testing for common trends. *Journal of the American Statistical Association* **83**: 1097-1107.
- Stock JH, Watson MW. 2002. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics* **20**: 147-162.
- Stott PA, Jones G, Lowe J, Thorne P, Durman C, Johns T, Thelen J. 2006. Transient climate simulations with the HadGEM1 model: causes of past warming and future climate change. *Journal of Climate* **19**: 2763-2782.
- Swanson KL, Sugihara G, Tsonis AA. 2009. Long-term natural variability and 20th century climate change. *Proceedings of the National Academy of Sciences* **106**: 16120-16123.
- Tol RSJ, De Vos AF. 1993. Greenhouse statistics-time series analysis. *Theoretical and Applied Climatology* **48**: 63-74.
- Tol RSJ, De Vos AF. 1998. A Bayesian statistical analysis of the enhanced greenhouse effect. *Climatic Change* **38**: 87-112.
- Tollefson J. 2014. Climate change: The case of the missing heat. *Nature* **505**: 276-278.
- Tollefson J. 2016. Global warming "hiatus" debate flares up again. *Nature* doi:10.1038/nature.2016.19414.
- Trenberth KE. 1984. Signal versus noise in the Southern Oscillation. *Monthly Weather Review* **112**: 326-332.
- Trenberth KE. 2015. Has there been a hiatus? *Science* **349**: 691-692.
- Trenberth KE, Fasullo JT. 2013. An apparent hiatus in global warming? *Earth's Future* **1**: 19-32.
- Triacca U. 2001. On the use of Granger causality to investigate the human influence on climate. *Theoretical and Applied Climatology* **69**: 137-138.

- Velders GJM, Andersen SO, Daniel JS, Fahey DW, McFarland M. 2007. The importance of the Montreal Protocol in protecting climate. *Proceedings of the National Academy of Sciences* **104**: 4814-4819.
- von Storch H, Zwiers FW. 1999. *Statistical analysis in climate research*. Cambridge University Press, Cambridge.
- von Storch H, Navarra A. 1999. *Analysis of climate variability: Applications of statistical techniques*. Springer Science & Business Media, Berlin.
- Wilks DS. 2011. *Statistical methods in the atmospheric sciences*. Academic Press, New York.
- Wolter K, Timlin MS. 1998. Measuring the strength of ENSO - how does 1997/98 rank? *Weather* **53**: 315-324.
- Wu Z, Huang NE, Wallace JM, Smoliak BV, Chen X. 2011. On the time-varying trend in global-mean surface temperature. *Climate dynamics* **37**: 759-773.
- Zhang Y, Wallace JM, Battisti DS. 1997. ENSO-like interdecadal variability: 1900-93. *Journal of Climate* **10**: 1004-1020.
- Zivot, E., Andrews, DWK. 1992. Further evidence on the great crash, the oil price shock and the unit root hypothesis. *Journal of Business and Economic Statistics* **10**: 251-270.
- Zwiers F, Hegerl G. 2008. Climate change: Attributing cause and effect. *Nature* **453**: 296-297.

Table 1. Tests for a unit root allowing for a one-time break in the trend function applied to G, NH, SH, WMGHG, TRF and PC1.

Series	T_B	W	k	$\hat{\mu}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$t_\alpha(\hat{\lambda}_{tr}^{AO})$
G ^N	1971 (1962, 1980)	14.19 ^a	0	-0.3982 (-12.89)	0.0035 (9.01)	0.0134 (11.00)	-5.45 ^a
G ^H	1978 (1972, 1984)	24.96 ^a	0	-0.6268 (-15.58)	0.0053 (10.86)	0.0202 (10.43)	-6.46 ^a
NH ^N	1982 (1974, 1990)	12.91 ^a	0	-0.4568 (-12.40)	0.0045 (10.41)	0.0203 (9.95)	-5.75 ^a
NH ^H	1982 (1977, 1987)	27.89 ^a	0	-0.6243 (-13.32)	0.0055 (9.96)	0.0281 (10.85)	-7.16 ^a
SH ^N (1)	1925 (1917, 1933)	18.70 ^a	0	-0.0222 (-0.46)	-0.0045 (-5.17)	0.0144 (12.60)	-4.79 ^a
SH ^N (2)	1959 (1947, 1971)	0.55	0	-0.3036 (-8.41)	0.0018 (3.72)	0.0115 (10.31)	-4.65 ^a
SH ^H	1976 (1966, 1986)	7.52 ^a	0	-0.6700 (-17.32)	0.0055 (11.57)	0.0102 (5.84)	-8.02 ^a
WMGHG	1960 (1959, 1961)	20.19 ^a	7	-0.2872 (-23.68)	0.0105 (64.05)	0.0349 (87.22)	-3.94 ^c
TRF	1960 (1956, 1964)	4.46 ^a	1	-0.2401 (-10.56)	0.0064 (20.89)	0.0218 (29.09)	-4.25 ^b
PC1 ^N	1962 (1957, 1967)	51.41 ^a	0	-1.1055 (-26.14)	0.0112 (14.01)	0.0403 (19.68)	-5.95 ^a
PC1 ^H	1968 (1964, 1972)	86.72 ^a	0	-1.1997 (-27.88)	0.0139 (18.10)	0.0403 (17.20)	-7.55 ^a
G ^B	1984 (1974, 1994)	6.40 ^a	0	-0.5962 (-20.38)	0.0056 (16.39)	0.0129 (7.36)	-5.99 ^a
G ^K	1972 (1962, 1982)	9.70 ^a	0	-0.6418 (-20.57)	0.0037 (9.56)	0.0129 (10.21)	-4.98 ^a

The regression for the unit root tests is defined in regression (4). The symbols used are defined as follows T_B is the estimate of the break date; W is Perron-Yabu structural change test statistic; k is the number of lagged differences added to correct for serial autocorrelation; μ , $\hat{\beta}_1$ and $\hat{\beta}_2$ are the regression coefficients of the trend function and the corresponding t-statistic values are shown in parenthesis. Bold numbers denote statistical significance at the 5% levels. $t_\alpha(\hat{\lambda}_{tr}^{AO})$ is the Kim-Perron (Kim and Perron, 2009) unit root test statistic. a, b, c denote statistical significance at the 1%, 2.5% and 5% levels, respectively.

Table 2a. Test for nonlinear co-trending around a linear trend amongst G, NH, SH, RFGHG and TRF.

r	NASA	HadCRUT4	10% critical región	5% critical región
1	0.0349	0.03922	>0.11962	>0.15099
2	0.06018	0.06477	>0.16918	>0.20264
3	0.07745	0.06781	>0.21407	>0.25221
4	0.16067	0.14257	>0.25132	>0.29475
5	0.37723	0.38562	>0.28287	>0.33094

Bold figures are significant at the 10% level.

Table 2b. Test for nonlinear co-trending around a constant amongst G, NH, SH, RFGHG and TRF.

r	NASA	HadCRUT4	10% critical region	5% critical region
1	0.0489	0.06363	>0.35183	>0.46577
2	0.07494	0.0663	>0.53561	>0.67420
3	0.09744	0.12934	>0.70366	>0.86038
4	0.18129	0.20914	>0.86182	>1.03454
5	1.58117	1.67226	>1.01416	>1.21948

Bold figures are significant at the 10% level.

Table 3. Tests for a second break in the slope of the trend function of temperature and radiative forcing series.

Series	T_B	W	Series	T_B	W
WMGHG	1994 (1990, 1998) [-25%]	3.42 ^a	PC1,7 ^H	2004 (1986, 2022)	-0.25
TRF	1991 (1989, 1993) [-56%]	18.21 ^a	PC1,8 ^H	1989 (1981, 1997) [-29%]	1.76 ^b
G ^N	2005 (1994, 2016)	-0.10	PC1,9 ^H	1989 (1982, 1996) [-31%]	2.54 ^b
G ^H	2007 (2002, 2012) [-56%]	3.93 ^a	PC1,10 ^H	1989 (1982, 1996) [-29%]	2.53 ^b
G ^B	2005 (2000, 2010) [-49%]	2.84 ^b	PC1,2 ^N	1989 (1985, 1993) [-38%]	7.56 ^a
G ^K	2004 (1991, 2017)	-0.10	PC1,3 ^N	1987 (1982, 1992) [-38%]	4.59 ^a
PC1 ^N	1989 (1985, 1993) [-40%]	9.92 ^a	PC1,4 ^N	1987 (1982, 1992) [-27%]	5.67 ^a
PC1 ^H	1991 (1984, 1998) [-31%]	2.53 ^b	PC1,5 ^N	1981 (1976, 1986)	0.26
PC1,2 ^H	1989 (1981, 1997) [-25%]	1.54 ^c	PC1,6 ^N	1989 (1985, 1993) [-47%]	9.99 ^a
PC1,3 ^H	2002 (1995, 2009)	0.10	PC1,7 ^N	1988 (1977, 1999)	0.11
PC1,4 ^H	1989 (1981, 1997) [-23%]	2.36 ^b	PC1,8 ^N	1989 (1981, 1997) [-26%]	2.13 ^b
PC1,5 ^H	2004 (1998, 2010)	0.98	PC1,9 ^N	1988 (1984, 1992) [-39%]	9.73 ^a
PC1,6 ^H	1989 (1983, 1995) [-33%]	2.86 ^b	PC1,10 ^N	1988 (1984, 1992) [-39%]	10.06 ^a

W is Perron-Yabu structural change test statistic. a, b, c denote statistical significance at the 1%, 5% and 10% levels, respectively. Figures in brackets denote the percent change in the slope coefficient after the second break.

Table 4a. Factor loadings of the rotated principal component analysis of NASA's G, NH, SH, and WMGHG, TRF, AMO, SOI, NAO, PDO and STRAT.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
GNASA	0.96	0.01	-0.06	0.08	-0.22	0.08	-0.09	0.11	0.00	-0.01
NHNASA	0.90	0.01	-0.13	0.06	-0.33	0.05	0.08	0.21	0.00	0.00
SHNASA	0.94	0.02	0.04	0.09	-0.08	0.11	-0.28	-0.03	0.00	0.00
WMGHG	0.98	0.05	0.02	0.09	0.04	-0.01	0.12	-0.08	-0.05	0.00
TRF	0.98	0.04	0.03	0.06	0.04	-0.01	0.10	-0.11	0.04	0.01
AMO	0.16	0.02	-0.11	0.15	-0.97	0.02	-0.01	0.01	0.00	0.00
SOI	-0.10	0.27	-0.09	0.08	0.02	-0.95	0.01	0.00	0.00	0.00
VOLCANO	0.02	0.08	-0.98	0.09	-0.11	-0.08	0.00	0.01	0.00	0.00
NAO	-0.16	0.02	0.10	-0.97	0.15	0.07	0.00	0.00	0.00	0.00
PDO	-0.07	-0.96	0.09	0.02	0.02	0.26	0.00	0.00	0.00	0.00

Extraction: principal components. Rotation: varimax normalized. Correlations higher than 0.70 are shown in bold.

Table 4b. Factor loadings of the rotated principal component analysis of HadCRUT4's G, NH, SH, and WMGHG, TRF, AMO, SOI, NAO, PDO and STRAT.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
GHADLEY	0.95	0.02	0.12	0.05	-0.22	0.04	0.17	-0.04	0.01	0.00
NHHADLEY	0.93	0.04	0.14	0.03	-0.23	0.00	0.23	0.06	0.01	0.00
SHHADLEY	0.92	-0.03	0.07	0.10	-0.16	0.13	0.00	-0.30	0.00	0.00
WMGHG	0.98	0.06	-0.06	0.11	0.05	0.00	-0.14	0.08	-0.05	0.00
TRF	0.97	0.05	-0.07	0.08	0.04	0.01	-0.18	0.09	0.04	0.00
AMO	0.18	0.02	0.11	0.15	-0.97	0.02	0.01	-0.01	0.00	0.00
SOI	-0.08	0.27	0.09	0.08	0.02	-0.95	0.00	0.01	0.00	0.00
VOLCANO	0.06	0.08	0.98	0.09	-0.11	-0.08	0.01	0.00	0.00	0.00
NAO	-0.15	0.02	-0.10	-0.97	0.15	0.07	0.00	0.00	0.00	0.00
PDO	-0.07	-0.96	-0.08	0.02	0.02	0.26	0.00	0.00	0.00	0.00

Extraction: principal components. Rotation: varimax normalized. Correlations higher than 0.70 are shown in bold.

Table 5a. Eigenvalues and percent of total variance accounted by the ith principal component. NASA's G, NH, SH, and WMGHG, TRF, AMO, SOI, NAO, PDO and STRAT.

	Eigenvalue	% Total variance	Cumulative %
1	4.88	48.82 (46.18)	48.82 (46.18)
2	1.71	17.12 (10.04)	65.94 (56.22)
3	1.29	12.92 (10.26)	78.86 (66.49)
4	0.80	7.96 (10.06)	86.82 (76.55)
5	0.71	7.10 (11.41)	93.91 (87.95)
6	0.43	4.27 (10.06)	98.18 (98.02)
7	0.11	1.14 (1.18)	99.32 (99.20)
8	0.06	0.63 (0.76)	99.96 (99.96)
9	0.00	0.04 (0.04)	100.00 (100.00)
10	0.00	0.00 (0.00)	100.00 (100.00)

Figures in parenthesis denote the percent of total variance accounted by the rotated ith principal component.

Table 5b. Eigenvalues and percent of total variance accounted by the ith principal component. HadCRUT4's G, NH, SH, and WMGHG, TRF, AMO, SOI, NAO, PDO and STRAT.

	Eigenvalue	% Total variance	Cumulative %
1	4.86	48.58 (45.88)	48.58 (45.88)
2	1.69	16.93 (10.06)	65.51 (55.94)
3	1.28	12.75 (10.45)	78.27 (66.39)
4	0.82	8.22 (10.12)	86.49 (76.51)
5	0.70	7.03 (10.99)	93.51 (87.50)
6	0.43	4.35 (10.07)	97.86 (97.57)
7	0.12	1.18 (1.33)	99.04 (98.90)
8	0.09	0.92 (1.06)	99.96 (99.96)
9	0.00	0.04 (0.04)	100.00 (100.00)
10	0.00	0.00 (0.00)	100.00 (100.00)

Figures in parenthesis denote the percent of total variance accounted by the rotated ith principal component.

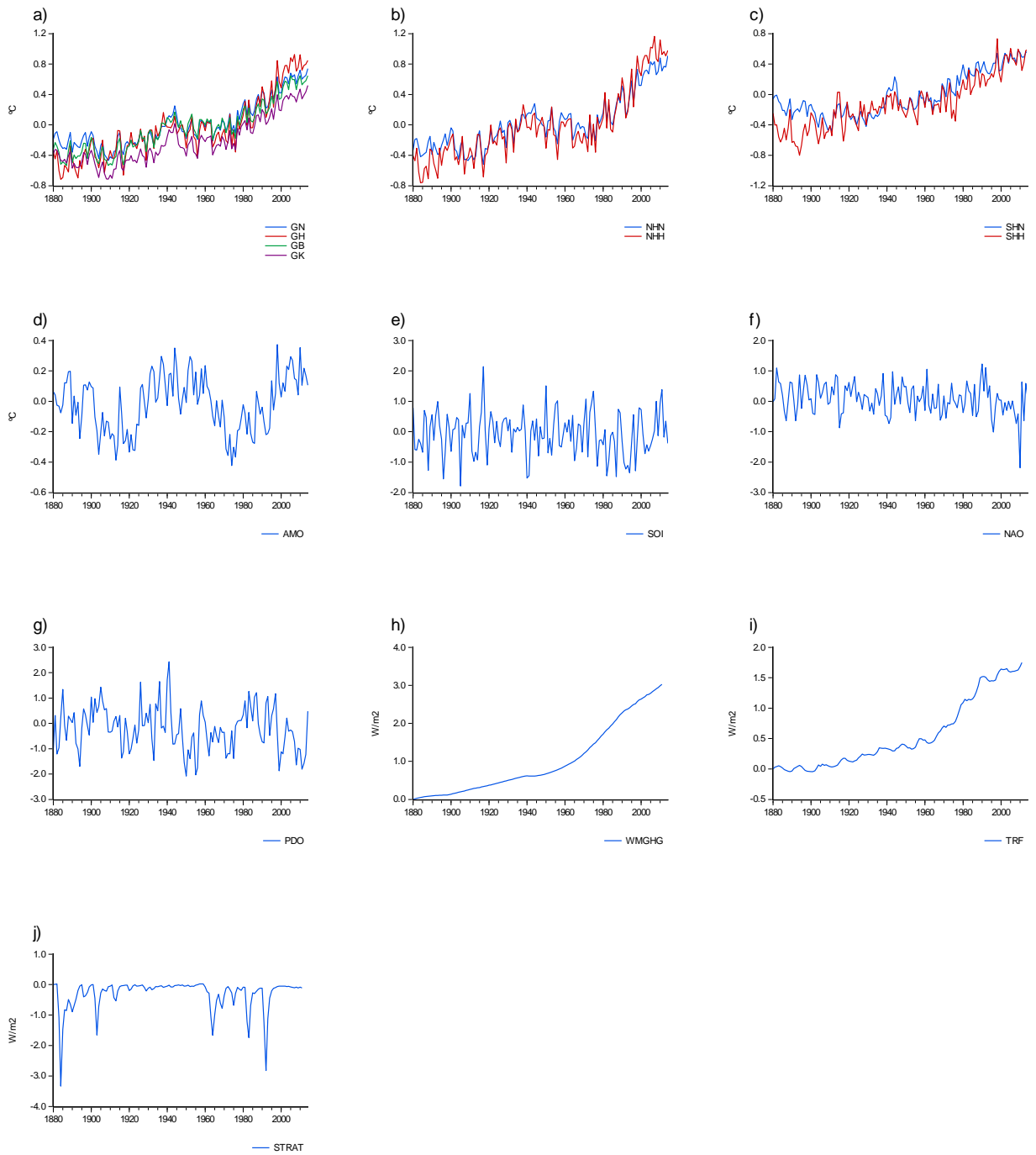


Figure 1. Global and hemispheric temperatures, natural variability modes and radiative forcing series. G, NH and SH temperature series are shown in panels a, b and c, respectively. AMO, SOI, NAO and PDO are shown in panels d, e, f and g, respectively. WMGHG, TRF and STRAT are shown in panels h, i and j.

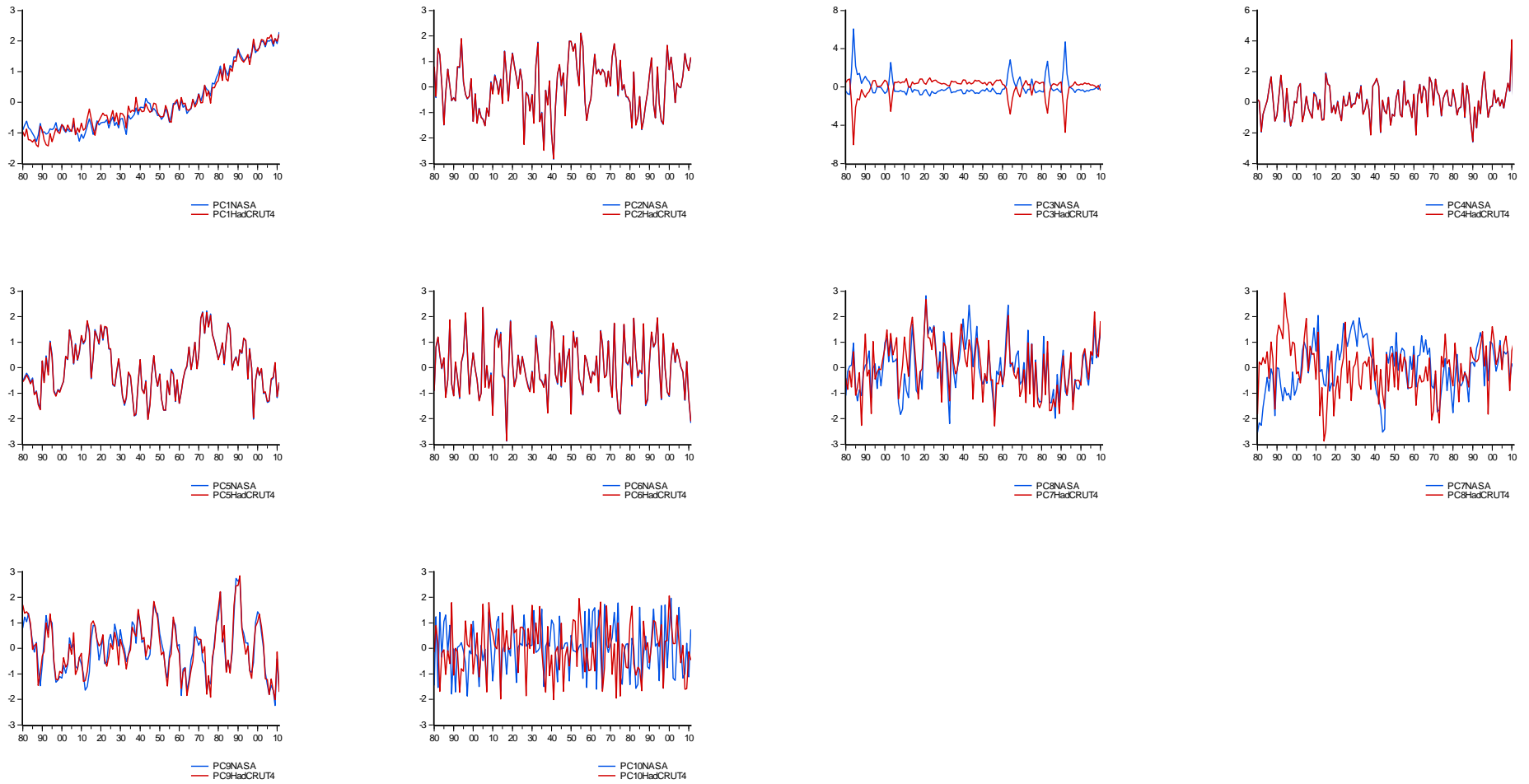


Figure 2. Rotated principal components of G, NH, SH, WMGHG, TRF, STRAT, AMO, SOI, NAO, PDO.