Breaks, trends and the attribution of climate change: a time-series analysis

Francisco Estrada*

Universidad Nacional Autónoma de México and Vrije Universiteit

Pierre Perron†

Boston University

March 20, 2012 (Preliminary Draft)

Abstract

Climate change detection and attribution have been the subject of intense research and debate over at least four decades. However, direct attribution of climate change to anthropogenic activities using observed climate and forcing variables through statistical methods has remained elusive, partly caused by the difficulties for correctly identifying the time-series properties of these variables and by the limited availability of methods for relating nonstationary variables. This paper provides strong evidence concerning the direct attribution of observed climate change to anthropogenic greenhouse gases emissions by first investigating the univariate time-series properties of observed global and hemispheric temperatures and forcing variables and then by proposing statistically adequate multivariate models. The results show that there is a clear anthropogenic fingerprint on both global and hemispheric temperatures. The signal of the well-mixed GHG forcing in all temperature series is very clear and accounts for most of their secular movement since the beginning of observations. Both temperature and forcing variables are characterized by piecewise linear trends with abrupt changes in their slopes estimated to occur at different dates. Nevertheless, their long-term movements are so closely related that the observed temperature and forcing trends cancel out. The warming experimented during the last century was mainly due to the increase in GHG which was partially offset by the effect of tropospheric aerosols. Other forcing sources, such as solar, are shown to only contribute to (shorter-term) variations around the GHG forcing trend.

*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

Both contributions of the Working Group I and of Working Group II of the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC-WGI, 2007; IPCC-WGII, 2007) exhibit an ever-growing bulk of direct and indirect scientific evidence regarding the warming of the climate system during the 20th century and of the role anthropogenic activities have played on it. The detection and attribution of climate change is no longer limited to changes in climate variables. Studies have shown that the warming signal is so strong that other physical and biological systems have been affected 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-WGII, 2007). An important part of the evidence on attribution is based on comparing observations to model predictions about what the state of a variety 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 change to externally forced climate change signals (IPCC-WGI, 2007). In this method, the dependent variable is usually an observed climate record and the independent variables 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 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-WGI, 2007, see also Stott et al., 2006b; Mitchell et al., 2001).

However, direct attribution of climate change to anthropogenic activities using observed climate and forcing variables through statistical methods has remained elusive. Important examples of the application of modern statistical techniques to address the attribution of climate change are the work of Stern and Kaufmann (1997a,b), Kaufmann et al. (2006, 2009), Tol and de Vos (1993;1998), among others. These articles provided 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 for addressing the detection and attribution of climate change (Gay et al., 2009; Estrada et al., 2010; Estrada et al., 2011; Mills 2010; Gil-Alana 2008). The recent literature has

shown that the assumption on which these earlier attribution studies are based on (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).

This paper first investigates the time-series properties of temperatures and forcing variables using state-of-the-art econometric techniques to determine what type of data generating process best describes them. Establishing these properties allows then to test for the existence of long-term relationships between temperatures and forcing variables using the class of multivariate models that are most adequate for dealing with the type of nonstationarities that are present in the data.

It is shown that both temperature and different combinations of the radiative forcing variables can be best represented as trend stationary processes with a one-time structural change in the slope of the trend functions occurring at different dates and of diverse magnitudes. In spite of these differences, the multivariate models indicate that the long-term movements of all these variables are so closely related that the observed temperature and forcing trends cancel out producing stationary residuals. The results show a clear anthropogenic fingerprint on both global and hemispheric temperatures. In particular, well-mixed GHG is the variable which contributes the most to the secular movement of the total observed radiative forcing and the signal of this variable is clearly imprinted in all temperature series and accounts for most of their long-term movement since the beginning of observations.

The analyses suggest that the break in the slope of the trend function of the first principal component (as well as of the partial and total sums) of the forcing variables leads to an abrupt change in the slope of the southern hemisphere temperature series about 8 years later and the occurrence of this response is delayed for about 37 years in the northern hemisphere, possibly because of complex ocean-atmosphere dynamic processes that are documented in the literature (Wu et al., 2011; Knudsen et al., 2010; Schlesinger and Ramankutty, 1994; Kerr, 2000). The post-break trend slope of the first principal component of forcing variables is two times larger than its first slope value and the resulting post-break trend slopes of global, northern and southern hemispheres temperatures are up to 7, 9 and 10 times larger than their pre-break slope values, respectively. The contribution of the different forcing variables is also estimated and it is shown that the warming experimented during the last century was mainly due to the increase in GHG which was partially offset by the effect of tropospheric aerosols. Other forcing sources, such as solar, mainly contribute to shorter-term variations around the GHG forcing trend. Nevertheless, these variations around the forcing trend are shown to play no role on the short-run dynamics of temperature series, at least in annual time-scales.

2 Univariate analysis of the time-series properties.

For more than two decades there has been a debate regarding the time series properties of global and hemispheric temperatures. 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 Stern and Kauffman (1997a,b) and Kauffman et al. (2006, 2009) proposed the use of cointegration techniques for representing the long-term relationship between temperature and forcing variables, providing at the same time strong evidence on the detection and attribution of climate change and seemingly ending the debate regarding the data generating process of temperature series. Nevertheless, all the results and inferences provided in such papers rely on the hypothesis of the presence of unit roots in temperature and forcing series, a proposition that was never soundly tested. Recently, the hypothesis of the existence unit roots in observed and simulated temperature series has been seriously questioned based on statistical and climate physics grounds (Gay et al., 2009; Estrada et al., 2010; Estrada et al., 2011).

2.1 Methodology.

The time-series properties of global and hemispheric temperatures from both the Climatic Research Unit's HadCRUT3 (Brohan et al., 2006) and the NASA database (Hansen et al., 2010) are analyzed [1]. These databases differ basically in three aspects: 1) the way temperatures are extrapolated (or not) into regions without observing stations. This is particularly important since the HadCRUT3 excludes most of the Arctic, where the warming has been very large during the past decade; 2) the ocean datasets used and; 3) the sampling period of the HadCRUT3 spans the period 1850-2010 while NASA covers only 1880-2010 (see Hansen et al., 2010). The series are presented in Figure 1. Below, we show strong evidence for TS when the recent state-of-the-art unit root tests that allow for a break in the trend function are applied (Kim and Perron, 2009; Perron 1997, 1989). These results are broadly similar to those of Gay et al. (2009) and Estrada et al. (2011), except that here we provide evidence for the presence of two breaks in the slope of the trend function of the southern hemisphere's temperature record. We also show these breaks to also be present in the radiative forcing produced by well-mixed greenhouse gases.

The standard Augmented Dickey-Fuller unit root test (ADF, Dickey and Fuller, 1979; Said and Dickey, 1984) applied to global and hemispheric temperatures seems to provide very strong evidence for DS processes (Table 1) [2]. However, as shown by Perron (1989) these results can be

[1] These are currently available at http://hadobs.metoffice.com/hadcrut3/diagnostics/index.html and http://data.giss.nasa.gov/gistemp.

[2] Only the results of the Augmented Dickey-Fuller test are shown here, but the results from other four commonly

misleading if there are breaks in the trend function of the time-series under investigation. In such cases, the sum of the first order autoregressive coefficients is highly biased towards unity, so that the unit root null is hardly rejected even if the series is composed of i.i.d. disturbances around a deterministic trend. Furthermore, if the break occurs in the slope of the trend function, unit root tests are not consistent; the null hypothesis of a unit root cannot be rejected even asymptotically.

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; WGI-IPCC, 2007; Gay et al., 2009). As mentioned above, standard unit root tests can be seriously affected by the presence of such breaks and the results in Table 1 could be misleading. To test for the existence of these breaks we applied the testing procedure of Perron and Yabu (2009) 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 for applying structural change tests. According to the Perron-Yabu (2009) test (Table 2), 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.

Formal testing for unit roots in the presence of structural breaks started with the paper of Perron (1989). He proposed three models that allow for the possibility of a one-time break in the intercept and/or slope of the trend function that is assumed to occur at a known date. The three models are: the "crash" model where a change in the intercept is permitted; the "changing growth" model that allows for a change in the slope; and a mixed model where a change in both intercept and slope occurs simultaneously. Several methodologies have been developed to extend his original test for the case of an unknown break date, the most widely used being the Perron (1997) and the Zivot and Andrews (1992) tests. For these tests the possibility of a break is only allowed under the alternative hypothesis (stationary noise component) and not under the null (integrated noise component). The recent literature has shown that these tests can have a major drawback, namely a rejection of the null can occur if there is a large break in the slope of the trend function albeit a unit root is present (see Vogelsang and Perron, 1998, and Kim and Perron, 2009).

To avoid this problem, Kim and Perron (2009) present a new generation of unit root testing procedures based on the original test of Perron (1989) that allow for a break in the trend function both under the null and alternative hypotheses and that include a pretest for the existence of a break using the Perron-Yabu (2009) test. If this pre-test rejects, the limit distribution of the

used unit root tests are available upon request. All of them strongly suggest the presence of a unit root in temperature series.

unit root test is then the same as if the break date was known (see Perron, 1989 and 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. For the interested readers, the models and tests are presented in more detail in the Appendix Methods.

2.2 Results for observed global and hemispheric temperatures.

We apply the Kim and Perron (2009) and the Perron (1997) tests to investigate the type of data generating processes of both temperature and radiative forcing variables (see the next section). For all of 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, as was also proposed in previous papers (Gay et al., 2009; Estrada et al., 2011). As shown in Table 3, the results of applying the Perron (1997) and Kim and Perron (2009) tests provide quite clear evidence for TS once a break in the slope of the trend function is allowed. For all time series and with both tests, the null hypothesis of a unit root is rejected at the 1% significance level in favor of a trend stationary process with one-time permanent change in the rate of growth. That is, 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 the global and hemispheric temperatures and have consisted in abrupt increases of different magnitudes in the rate of warming.

It is important to note that in the case of global an northern hemisphere temperatures the estimates of the break dates are quite similar for both the HadCRUT3 and the NASA time series: 1971 and 1978 and 1984 and 1985 for the global and northern hemisphere temperatures from the HadCRUT3 and NASA, respectively. As revealed by the results of applying the methodology of Perron and Zhu (2005), when considering a 95% confidence interval the break dates estimated from these databases cannot be considered statistically different (Table 4). It is also worth noting that the break dates from global and northern hemisphere temperatures are not statistically different, regardless of which database is used. These results are also congruent with those shown in Gay et al. (2009). Nevertheless, the estimates of the break date in the southern hemisphere temperature are quite different, and depend on the database selected. The results obtained using the NASA dataset are very similar to those in Gay et al. (2009), where the break in the trend function is estimated to occur at the beginning of the 20th century. On the contrary, using the HadCRUT3 the break date is estimated to occur in 1955, a break date that is statistically different from both the estimate using the NASA dataset and that of Gay et al. (2009).

These differences can be explained by analyzing the sum of the squared residuals (SSR) of the regressions used for estimating the beak date. The estimates of the SSR show two minima in both the NASA and HadCRUT3 series. Figure 2 reveals that the results are highly sensitive to the sample period used: when the sample of the HadCRUT3 series is restricted to the same sample period as that of the NASA series (Figure 2b), the global minimum occurs around 1915 and is very similar to that of NASA (Figure 2a). However, when the full sample of the HadCRUT3 is used (1850-2010) the global and local minima exchange places: the global minimum occurs in 1955 and the local one takes place during the first three decades of the century (Figure 2c). This behavior suggests the existence of a second break in the slope of the trend function and that the longer sample of the HadCRUT3 provides a better estimate of the break date. For this reason, Table 3 also includes the results for the NASA southern hemisphere temperature series when choosing 1955 as the break date (labeled SH_N*). As can be seen form such table, the rejection of the null of a unit root is robust to this change in the break date.

The total increase in global and hemispheric temperatures is estimated to be practically the same from the HadCRUT3 and NASA datasets: near 0.8^oC, 1^oC and 0.7^oC, for the global, northern and southern hemisphere temperatures. Moreover, according to both datasets about 60% of the observed warming in the globe and northern hemisphere has occurred in a very short period of time: during the last three to four decades in the case of the global average and in the last twenty-five years for the northern hemisphere. The rate of increase in the global and northern hemisphere temperatures in the post-break period has been approximately 1.7^oC and 2.7^oC per century, respectively. In the case of the southern hemisphere, most of the warming has occurred since the last 55 years (between 64% and 82%, depending on the dataset) and the rate of increase in temperature has been approximately 1^oC per century.

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. In the case of global and northern hemisphere temperatures, the estimates of the pre-break slope using NASA's dataset are 77% and 40% larger than those using the HadCRUT3 dataset, and 145% larger in the case of the southern hemisphere temperature. The differences in the estimated post-break slopes are small in the cases of global and northern hemisphere temperatures (5% and 1%, respectively) but for the NASA dataset the southern hemisphere post-break slope is about 34% smaller than that of the HadCRUT3's. The large differences in the estimates of the southern hemisphere pre- and post-break slopes using NASA and HadCRUT3 datasets can be due to the inclusion of a wider coverage in the sea-surface temperature dataset in the later (see Brohan, 2006).

According to these estimates observed temperatures have experienced a large increase in the

rate of growth since the middle of the 20th century. For the NASA dataset the post-break rate of warming (i.e., the pre-break slope plus post-break slope) is 3.3, 5.6 and 2.4 times larger than the pre-break estimate for global, northern and southern hemisphere, respectively. For the HadCRUT3 dataset these rates are 6.2, 8.1 and 9 times larger than the pre-break slope, in that order. As will be shown in the next sections, these abrupt increases represent the response of the climate system to a doubling in the rate of growth of the observed forcing.

These changes concur with the important characteristics of observed warming that have been previously reported (Gay et al., 2009): temperature series can be better described as trend stationary processes with a one-time permanent shock which cannot be interpreted as part of the natural variability; climate change has occurred in two stages, the first consisting in a moderate rate of warming and a second one showing a much larger rate of increase occurring rather abruptly in a short period of time. Recently Estrada et al. (2011) have shown that these time-series characteristics are also present in the 20th Century Climate Experiment (20c3m) simulations conducted for the IPCC's Fourth Assessment Report (IPCC-WGI, 2007).

2.3 Results for the observed radiative forcing.

While there has been a debate about the time-series properties of global and hemispheric temperatures, radiative forcing variables have received little attention in this respect and have been usually assumed to be integrated processes when conducting attribution studies based on observed records and time series analysis (Stern and Kaufmann, 1997ab; Kaufmann et al., 2006, 2009; Tol and de Vos, 1993, 1998). In this section, the time series properties of radiative forcing series are investigated using the same econometric techniques applied before to temperature series. The results are robust to different definitions of the radiative forcing trend and show that the common assumption of forcing variables being integrated processes is not supported by the data.

The radiative forcing series used in this section were published by Hansen et al. (2011) and are available at the NASA Goddard Institute for Space Studies website (http://data.giss.nasa.gov/modelforce/RadF.txt). The database covers the period 1880-2010 and includes the following variables (in W/m2): well mixed greenhouse gases (carbon dioxide (CO2), methane (NH4), nitrous oxide (N2O) and chlorofluorocarbons (CFCs)); ozone (O3); stratospheric water vapor (H2O); solar irradiance; land use change; snow albedo; stratospheric aerosols; black carbon; reflective tropospheric aerosols and; the indirect effect of aerosols. As can be seen from Figure 3, with the exception of stratospheric aerosols, all time series show a clearly nonstationary behavior in their level with large changes in their rates of growth.

Because the interest relies on investigating the relationship between temperature and anthro-

pogenic and natural forcing series, different "forcing trends" that may be driving the observed changes in global and hemispheric temperatures are constructed and analyzed in this section, instead of analyzing the time-series properties of the individual forcing variables. In addition, using these radiative forcing trends for estimating the multivariate models shown in the next section, as a substitute for the original forcing variables, permits avoiding severe multicollinearity problems since the variables in Figure 3 typically show correlations close to 0.99. Two approaches were adopted for constructing different "forcing trends".

The first approach consists in simply adding the forcing variables to construct the forcing trend. This is the standard procedure in climate modeling. Three forcing trends using this approach were considered (Figure 4): 1) well-mixed greenhouse gases (WM_GHG) which represents the dominant human contribution to the observed radiative forcing; 2) GHG_SOLAR, composed of WM_GHG plus SOLAR, represents the main anthropogenic and natural forcing variables; 3) ALL_FORCING includes all of the human and natural forcing that contribute to the forcing trend (GHG_SOLAR plus H2O, LAND_USE, S_ALBEDO, BC, R_AER and AIE). Note that STRAT_AER was not added for constructing the forcing trend given that it is stationary around a constant (see below for the results of the unit root tests). Instead, this variable is introduced as another independent variable in the regression models that are presented later. As can be seen from Figure 4, WM_GHG is the most important variable contributing to the increase of radiative forcing. Solar forcing mainly contributes to adding variability to the WM_GHG trend, particularly imparting the well-known Schwabe 11 year cycle (WGI-IPCC, 2007). The ALL_FORCING variable shows that the sum of forcing factors other than greenhouse gases and solar have a negative net effect on the forcing trend. The reduction in the rate of growth of the total radiative forcing is particularly important since the 1990's, which is primarily due to the direct and indirect effects of tropospheric aerosols [???].

The second approach consists in applying a principal component analysis for reducing the dimensionality and extracting the principal modes of variation, in particular the trend mode. The first principal component explains 83.32% of the variance of the radiative forcing and, as can be seen in Figure 5, it represents the forcing trend. As revealed by the eigenvector associated to this principal component, the forcing trend is a mixture of human and natural factors with similar weights, except for stratospheric aerosols whose weight is close to zero. The forcing trend, presented in Figure 5, is thus a combination of all forcing series excluding stratospheric aerosols. The second principal component (also depicted in Figure 5) explains only 10.15% of the variance and it mainly represents the stratospheric aerosols forcing. The remaining principal components contribute only marginally to explaining the variance of this set of time series and therefore are excluded from subsequent analyses.

The time series properties of the series constructed using both approaches are analyzed in the rest of this section and will be used for the multivariate models presented in the following sections. Note that while the first approach has the advantage that it has a clear physical meaning and that it allows for easier attribution of the observed warming to different natural and anthropogenic factors, the second approach may provide a clearer separation of the trend mode from the other modes of variability.

As in the previous section, standard ADF unit root tests were applied to the forcing variables. The ADF tests included in all cases a constant and a trend as the deterministic component with the exception of the STRAT_AER and PC2 variables for which only a constant was included. The results in Table 6 show that while PC2 and STRAT_AER can be considered as stationary around a constant, the unit root null hypothesis cannot be rejected at any conventional levels for all the "forcing trend" variables. Nevertheless, as mentioned in the previous section, standard unit root tests can be severely affected by the presence of structural breaks in the trend function, particularly when the break occurs in the slope parameter. Figures 4 and 5 strongly suggest that the rate of growth of these variables has not remained constant over the sample period, and thus standard unit root tests may not be adequate for investigating their time-series properties.

The existence of structural breaks in the slopes of the trend function of these forcing trend variables was tested using the Perron and Yabu (2009) procedure and, as shown in Table 7, for all variables the null of parameter stability can be strongly rejected. Although the presence of structural breaks in the forcing variables has been discussed in the literature, it has seldom been formally tested (see for example, WGIII-IPCC, 2007). The results of the Perron-Yabu (2009) procedure clearly suggest that the finding of unit roots in the forcing variables reported in previous publications may due to the effect that breaks in the slope of the trend function have on standard unit root tests. As in the case of global and hemispheric temperatures, allowing for the presence of a one-time structural change in the slope of the trend function changes the conclusions markedly. As shown in Table 8, the Perron (1997) and Kim and Perron (2009) tests applied to the forcing trend variables defined above strongly reject the null of a unit root in favor of trend stationary processes with one permanent break in their rate of growth.

This finding has important implications for the multivariate modeling of temperatures and radiative forcing variables. The results indicate that there are no differences in the order of integration of these variables, all of them can better be described as trend stationary processes. Therefore, the existence of long-term relationships can be tested using standard regression models and there is no need to invoke for this purpose cointegration methods that require the assumption of unit root processes in temperature and forcing series. As mentioned, this assumption has been shown to be

at odds with the observed and simulated temperature data and the physics of climate (Gay et al., 2009; Estrada et al., 2010; Estrada et al., 2011; Triacca, 2001).

The point estimate of the break date for all the forcing trend variables constructed under the first approach is 1960. This break occurs after the small decline in the well-mixed greenhouse gases forcing around WWII and the subsequent recuperation of European economies (Figure 4, WM_GHG). For PC1, the break date occurs 13 years earlier, in 1947. The estimated post-break rates of growth of the radiative forcing variables defined above are of much larger magnitude in comparison with those before the pre-break: for WM_GHG the increase in the rate of growth is 334%; 213% for GHG_SOLAR and 345% for ALL_FORCING. In the case of PC1, the post-break rate of growth is 107% larger than it was prior to the occurrence of the break. It is interesting to note that according to the 95% confidence intervals estimated using the Perron and Zhu (2005) procedure (Table 9), the break dates of the forcing variables are not statistically different form the break date found in the southern hemisphere temperature series. Also, when applying the Perron and Yabu (2009) test on the forcing trends to the sub-sample 1880-1930, another break in the slope can be found at the beginning of the 20th century which is very close to the one that can be found in the southern hemisphere temperature series using the NASA dataset and when restricting the HadCRUT series to 1880-2010 (Table 4 and Figure 2). In the case of global and northern hemisphere temperatures, the break in the slope occurs later in the century and there is no evidence of another break at the beginning of the century. Despite these differences, it is shown in the next section that the long-term movements of all these variables are so closely related that the observed temperature and forcing trends cancel out producing stationary residuals. Temperature and forcing variables share the same long-term trend. Furthermore, it becomes apparent that the forcing signal of well-mixed greenhouse gases is present in all temperature series, being especially pronounced for southern hemisphere temperatures. The signal is much clearer there because it is dominantly covered by ocean which is known to modulate temperature variability much more than in the northern hemisphere, making it more stable around the long-term trend (Jones et al, 1986a,b; Hansen and Lebedeff, 1987).

The difference in the estimates of the break dates in the Northern hemisphere, Southern hemisphere and global temperatures can be accounted for by the effect of the Atlantic Multidecadal Oscillation (AMO) which represents ocean-atmosphere processes naturally occurring in the North Atlantic with a large influence over the Northern Hemisphere and global climates. As shown in Estrada et al. (2012), when this effect is filtered out the break dates for Global and Northern Hemisphere from the HadCRUT3 dataset are estimated to occur in 1955 and 1959, respectively. The NASA dataset leads to similar estimates of the break dates, namely 1956 and 1968. Hence,

once the AMO is accounted for, all temperature series have nearly the same break dates, which are very close to those of the forcing variables. Since the AMO is a stationary process, though with long swings, it is inconsequential whether one accounts for it when assessing the time series properties of the temperature series (existence of breaks, integration order, co-movements, etc.). It is, however, important to include it to account for short-term variations, and we will do so below when providing multivariate time series models for Northern and global temperatures.

3 Common breaking trends.

Since Engle and Granger (1987), cointegration techniques have become the standard procedure for testing and estimating long-term relationships between nonstationary variables. The attribution studies based on time series analyses that have been published are based on cointegration techniques, although the basic assumption needed for this procedure, namely variables that are integrated processes, was never soundly tested for temperature variables until recently (Gay et al., 2009; Estrada et al., 2011). The present paper extends these previous studies applying statistical methods that are appropriate for non-integrated.

It is important to bear in mind that a cointegration analysis is only one possibility for relating trends of non-stationary variables. For example, Engle and Kozicki (1993) showed that relationships between non-stationary variables can be established when linear combinations of different time series cancel out some "common features" such as trends and co-breaking in trend parameters. We use a simple approach for testing for common trends in temperature and forcing variables that consists in estimating an Ordinary Least Squares (OLS) regression and conducting a standard ADF test on the residuals. This is similar to the Engle-Granger two-steps cointegration test but for trend-stationary variables and therefore the relevant critical values are those tabulated for the standard ADF test with no deterministic terms included. The basic idea is to detrend global and hemispheric temperatures with the forcing trends defined previously and test if there are non-stationarities remaining in the residuals. If the residuals are stationary, it can be concluded that both temperatures and radiative forcing share the same long-term trend and that the observed warming during the 20th century can be attributed to anthropogenic activities.

The first step for testing for a common trend in temperature and radiative forcing variables consists in estimating an OLS regression model of the type:

$$T_{i,j,t} = \alpha + \beta FT_{r,t} + \varepsilon_t \tag{1}$$

where $T_{i,j,t}$ are the observed average temperatures for the global, northern and southern hemisphere ($i = 1, 2, 3$) and for the NASA and HadCRUT3 datasets ($j = 1, 2$), while $FT_{r,t}$ are the forcing

trends. The way forcing trends were defined above is useful to separate the importance of human and natural factors in the observed warming trend: WM_GHG can be assumed to be mostly human-made, while GHG_SOLAR represent a combination of the main human and natural forcing factors. ALL_FORCING represents the sum of all natural and anthropogenic forcing factors.

If the residuals of regression (1) obtained using WM_GHG are found to be non-stationary but to be stationary when using the GHG_SOLAR or ALL_FORCING variables, then this would be evidence about the strong contribution to observed warming by non-human-induced forcing factors. On the opposite, if the residuals are stationary when temperatures are detrended using WM_GHG only, the interpretation is then that anthropogenic factors are the dominant cause of global warming. To further highlight the importance of human role in the observed warming of the 20th century, a combination of the forcing trends defined above is also considered: ALL_FORCING minus WM_GHG (hereafter ALL_FORCING*) which represents all forcing with the exclusion of the most important anthropogenic forcing (i.e., greenhouse gases).

Tables 10 shows the slope coefficient estimates of regression (1) as well as the associated t-statistic value and the coefficient of determination for each combination of temperature and forcing trends. The first thing that transpires from these results is that the slope coefficients in all the estimated regressions using ALL_FORCING* are nonsensical because they imply a negative relation between radiative forcing and temperature that is not physically possible. The total forcing trend when excluding WM_GHG has a strong negative trend that cannot reproduce the observed warming. Hence, the observed warming cannot be explained in a physically consistent manner if the well-mixed greenhouse gases are excluded. The last column of Table 10 presents the value of the standard ADF test specified with no deterministic terms when applied to the residuals of regressions (1). These results show strong support for the common trend hypothesis between temperatures and radiative forcing, and in particular for the contribution of greenhouse gases to the observed warming. Even when the forcing trend is limited to WM_GHG, the ADF test is considerably larger (in absolute value) than the critical value at the 1% significance level, providing strong evidence that observed temperatures follow the same long-term path as the human-made forcing, which has been increasing very rapidly since 1960. Strong evidence in favor of a common trend is shown for global and hemispheric temperatures and for both NASA and HadCRUT3 datasets.

The variable PC1 is of special interest because, as mentioned, a principal component analysis permits extracting the trend mode in a way that it is more clearly separated from other modes of variation in the original multivariate forcing dataset. Hence, it may provide a better representation of the secular movement of the forcing during the sample period. The results of this procedure for testing for a common trend indicate a strong long-term relationship between temperature and

radiative forcing. It is worth noting that the common trend finding is not sensitive to the way the forcing trend is constructed as revealed by the ADF statistic values reported which are highly significant in all cases, provided WM_GHG is included.

Figure 6 shows the observed global and hemispheric temperatures from the NASA and HadCRUT3 datasets and the corresponding fitted temperatures obtained from regression (1). This figure illustrates that the different definitions of the forcing trend describe very well the long-term movement of all temperature series, irrespective of the temperature dataset used and of the forcing trend considered. It is worth noting that the fit produced by WM_GHG reproduces the southern hemisphere temperature so closely that it almost appears to be a running-mean of temperature. The fit is also very good for global northern hemisphere temperatures, though not as good as for the southern hemisphere. As has been documented in the literature, the northern hemisphere's temperature shows larger variability around its long-term trend because of large ocean-atmosphere cycles that are part of the natural variability over the North Atlantic (notably by the Atlantic Multidecadal Oscillation) and also by the differences in the percentage of land-continent (Wu et al., 2011; Knudsen et al., 2010; Schlesinger and Ramankutty, 1994; Kerr, 2000). The southern hemisphere temperature series depicts more clearly the secular movement than the global and northern hemisphere temperature because the ocean modulates much more temperature variability.

Figure 6 also shows the ALL_FORCING plus START_AER and the PC1 plus the PC2 fits [3], and as can be seen, these variables associated with volcanic eruptions have an important effect on temperature variability over the northern hemisphere that is also present in global temperatures. Stratospheric aerosols were not significant for the southern hemisphere temperature, showing that volcanic eruptions did not impacted significantly temperature in this hemisphere during the last century. Figure 6 illustrates other important common characteristics of temperatures and forcing trends, most of them being more evident for the southern hemisphere temperatures for the reasons mentioned. For example, the cooling or slow down of warming shown by all temperature series during the 1940-1975 period that has been reported numerous times in the literature (see for example, Jones et al., 1986a,b,c; Hansen and Lebedeff, 1987;Tol and de Vos, 1998) also coincides with a slow down in the well-mixed greenhouse gases forcing during the same period, which can be attributed to the ending of WWII and the subsequent recuperation of Europe's economy and industrial production. The relative slow down in the warming during the last two decades, which has sometimes been reported in the literature (Hansen et al., 2011), also matches with the deceleration in the increase of the well-mixed greenhouse gases forcing mainly in CFCs due to international

[3] For the global temperature series of the HadCRUT3 dataset, PC2 was not significant and therefore is not included in Figure 6.

regulation of these gases (this is more clearly seen in the southern hemisphere temperatures). Finally, as discussed in the previous sections, the structural break found in the NASA's southern hemisphere temperature (and in the HadCRUT3 data when the sample is restricted) is also present in the well-mixed greenhouse forcing.

The evidence presented in this section indicate that the WM_GHG forcing had a determinant role in the observed warming during the last century, and that is has been the main driver of the long-term movement in observed temperatures. The changes in aerosols, albedo and other forcing factors contributed somewhat to reducing the warming particularly since the second part of the 20th century. The solar cycles and the volcanic forcing are the main contributors to the variability around the well-mixed greenhouse forcing trend. Figure 7 illustrates these facts by showing the individual contribution of well-mixed greenhouse gases, solar and solar plus other forcing factors to the observed northern temperature series (NASA). The contributions of these variables to the observed warming were obtained by taking the differences between the fitted temperatures shown in Figure 6. The results are similar for global and southern hemisphere temperatures and therefore are not shown. While the WM_GHG have produced an increase of 0.85^oC in the northern hemisphere temperatures over the period of 1880-2010, the solar forcing produced an increase in temperatures around 0.06^oC from 1880 to 1960 and then a cooling of more or less the same magnitude from 1961 to 2010. The joint contribution of all forcing except WM_GHG amounts to a decrease of around 0.1^oC from 1880 to 2010. That is, the net contribution of all forcing factors other than WM_GHG has been a slight cooling of around 10% of the warming caused by the WM_GHG increase.

In the light of the strong evidence for a long-term relation between forcing and temperature variables, the results indicate that the break in the forcing trend caused the subsequent breaks in temperature series. Table 12 presents a comparison of the post- and pre-break values of the slope of the trend for the forcing and temperature series. In all cases the breaks in temperature series are more than proportional to the break in the forcing variable, irrespective of the forcing trend and of the temperature dataset. For example, for the case of PC1, a doubling in the rate of growth of the forcing produced a post-break slope in the global mean temperature between 4.3 (NASA) and 7.2 (HadCRUT3) times its previous value. For the northern hemisphere the response to the break in the forcing trend is even larger ranging from 6.6 (NASA) and 9.1 (HadCUT3) times its pre-break slope, while for the southern hemisphere the post-break slope is between 3.4 and 10 times its first slope value. Due to natural variability processes and climate dynamics, the breaks in the slope of the forcing manifested themselves with different time lags in global and hemispheric temperatures. The break had an impact on the southern hemisphere almost simultaneously, while for the northern hemisphere and global series these breaks occurred about 35 and 24 years later, respectively.

4 Short-run dynamics and Granger-causality tests.

To test for the existence of feedback processes in the short-run dynamics of forcing and temperature variables, a vector autoregressive model (VAR) was estimated using the detrended data of the PC1 forcing and of the NASA's global and hemispheric temperatures, as well as the Southern Oscillation Index (SOI) as endogenous variables and PC2 and a constant term as exogenous. The best VAR models estimated for both hemisphere and global temperature series are of order 3 and misspecification testing shows that they are statistically adequate (no autocorrelation, no heteroskedasticity, and residuals approximately normal). These models also satisfy the stability condition as the inverse roots of the characteristic polynomials lie inside the unit circle in all cases.

Table 12 presents the results of Granger Causality tests, which allow investigating the possible exogeneity of variables included in the VAR model. They show that the detrended forcing (PC1) does not Granger-cause detrended global or hemispheric temperatures, and that the detrended global or hemispheric temperatures also do not Granger-cause detrended PC1. Hence, the forcing trend, here represented by PC1, has a strong role defining the long-term evolution of global and hemispheric temperatures but does not contribute to the short-term dynamics and no significant short-term feedback processes is present, at least with the annual frequency considered. The results are the same using any of the forcing trends considered (as long as WM-GHG is included). Hence, the forcing trends are exogenous.

5 Time series models for global and hemispheric temperature series.

In this section, statistically adequate models for global and hemispheric temperatures from the NASA and HadCRUT3 datasets are presented. It is shown that parsimonious time-series models can provide a very good fit for observed global and hemispheric temperature series and that they produce basic estimates of climate sensitivity that are within the ranges that have been reported in the literature produced using different methods. The models attempt to capture both the short and long-term dynamics: the secular movement in global and hemispheric temperature is captured by the forcing trends that were shown to share their long-term trend with temperature variables, while the short-run dynamics are accounted for by the internal dynamics of temperature series and by some climate indices that are known to have an important role in capturing the inter-annual variability of global and hemispheric temperatures. The general form of the models is represented by the following regression:

$$T_t = \alpha + \beta_1 FT_t + \beta_2 AER_t + \sum_{j=1}^{J} \phi_j T_{t-j} + \sum_{k=0}^{K} \gamma_{k+1} SOI_{t-k} + \sum_{l=0}^{L} \delta_{l+1} NAO_{t-l} + \sum_{m=0}^{M} \eta_{m+1} AMO_{t-m} + \varepsilon_t \quad (2)$$

where T_t is global, northern or southern hemisphere temperature from the NASA and HadCRUT3 datasets. Two sets of models were estimated using two different forcing trends and stratospheric forcing. The first set uses the PC1 as the forcing trend variable (FT) and PC2 as a proxy for the stratospheric aerosols forcing (AER). The second set of models uses ALL_FORCING and STRAT_AER as the forcing trend and the stratospheric aerosols forcing, respectively. For modeling the short- and medium-term dynamics, the general model includes the El Niño/Southern Oscillation (ENSO, denoted here by the SOI), the North Atlantic Oscillation (NAO) and the Atlantic Multidecadal Oscillation (AMO), which are commonly considered as some of the most important natural sources of inter-annual global and hemispheric climate variability (Enfield et al., 2001; Hurrell, 1995; Wolter and Timlin, 1998).

Table 13 shows the estimated coefficients of the proposed models for global and hemispheric temperatures. Although the results of misspecification testing are not shown here, all models are statistically adequate in the sense that the residuals do not show autocorrelation, they are homoskedastic, normal and the estimated parameters are stable. Note that all coefficients have signs expected from climate physics: the forcing trend has a warming effect, while stratospheric aerosols produce a decrease in temperatures; a negative variation in SOI (for example, the occurrence of an El Niño) produces an increase in temperatures and the contemporaneous values of NAO of AMO are positively related to temperatures. The coefficient of determination R^2 in all of the models is very large, indicating that they fit observed data quite well.

The different models show strong similarities that make results robust in that they do not depend on the data set used (NASA or HadCRUT3) or the definitions of the forcing variables included (PC1 or ALL_FORCING; PC2 or STRAT_AER). Global temperature models typically include as significant variables the forcing trend, stratospheric aerosols, a lagged value of global temperature, and the contemporaneous values of SOI and AMO. As expected from the results in the previous sections, the short-term dynamics are more complex in the northern hemisphere temperatures compared to those for the global and southern hemisphere temperatures. Models for the northern hemisphere temperature in general include the same variables as for the global temperature models plus a second order lagged value of northern hemisphere temperatures and of AMO and a contemporaneous value of NAO. On the other hand, the models for the southern hemisphere temperatures show simpler short-term dynamics for which only the contemporaneous value of SOI and AMO and a lagged value of the southern hemisphere temperature are relevant for explaining short-term temperature variability. As discussed previously, the variability of the southern hemisphere temperatures is strongly modulated by the oceans and, hence, less affected by shorter-term processes compared to the global and northern hemisphere temperatures. As in

the case of the VAR model presented in the previous section, for all models of temperature series lagged terms of the forcing trend are not significant, supporting the hypothesis that this variable does not have a significant effect on short-term dynamics. Some differences in the results between the datasets are clearer for the southern hemisphere series, where for the HadCRUT3 series the forcing from stratospheric aerosols is not significant and the AMO is shown to be relevant for explaining part of the variability of temperature. These are due to the wider ocean sea-surface temperature coverage of the HadCRUT3 dataset (Brohan, 2006).

The estimated long-run effect of the total forcing (excluding the stratospheric aerosols) is very similar regardless of the dataset used: the global temperature estimate differs only by 1.08% from one dataset to the other, while for the northern hemisphere the difference is 5.86% and the largest difference (8.11%) occurs in the estimate for the southern hemisphere, where the two datasets are less similar. The estimates show even smaller differences amounting to less that 3.5% (0.15% for global, 3.12% and 3.44% for northern and southern hemisphere, respectively). Again, the results are robust to the different datasets and the definition of the forcing variables.

The estimates in Table 13 can provide basic approximation of the climate sensitivity (CS) to changes in the forcing variables. For example, a volcanic eruption of similar magnitude to that of Mount Pinatubo in 1991 would produce a cooling effect of 0.16^oC in global temperatures, and of 0.29^oC in the northern hemisphere. These values are close to those estimated from observations and climate models (Hansen, 1996) [4]. Basic estimates of CS to a doubling of the atmospheric concentrations of CO2 can also be obtained, although they must be interpreted with care. Strictly, to estimate CS two equilibrium states should be compared, and the observed climate is not at an equilibrium state (WGI-IPCC, 2007; Hansen, 2005). Rather, observed temperature records describe a climate transient, i.e., a measure of the strength and of the climate response to greenhouse gas forcing. Hence, the estimated CS presented here are expected to underestimate the "true" value of this parameter, and to lie close to the lower end of the CS range produced by climate models and reported, for example, in the IPCC's Fourth Assessment Report (WGI-IPCC, 2007).

To construct the CS estimates, it is assumed that for a doubling of CO2 atmospheric concentrations the corresponding radiative forcing is $\Delta Q_{2x} = 4.39$ W/m2 (Houghton et al., 1990; Wigley and Raper, 1990). The CS parameter is then estimated by $S = \beta_1/(1-\phi_1-\phi_2)\Delta Q_{2x}$. In order to account for the uncertainty in the estimated parameters β_1, ϕ_1 and ϕ_2, the confidence intervals for S where obtained from Monte Carlo simulations using 10,000 replications based on the parameters and standard errors shown in Table 13.

[4] These values correspond to the coefficients estimated using the NASA dataset. The estimates obtained using the HadCRUT3 dataset are similar: a cooling of 0.1 and 0.18 for global and northern hemisphere temperatures, respectively.

Table 14 presents the estimated CS for global and hemispheric temperatures for both the HadCRUT3 and NASA datasets. Note that the CS values are very similar for global and hemispheric temperatures irrespective of the dataset used. As expected, although the estimated mean CS values for global temperature (around 1.7^oC with a 95% confidence interval of 1.21^oC to 2.32^oC) are consistent with the range of CS estimates using climate models simulations (WGI-IPCC, 2007; Cubash and Meehl, 2001; Forests et al., 2002; Schlesinger, 2001, among others), they lie on the low end of such range. The estimates are closer to those produced using observations and statistical models (see, e.g., Kauffman et al., 2006; Stern and Kaufmann, 2000; Tol and de Vos, 1998).

The largest response to a doubling of CO2 atmospheric concentrations occurs in the northern hemisphere where the estimated 95% confidence interval extends from 1.17^oC to 3.14^oC, with a mean estimate of about 2^oC. The lowest values for CS are those for the southern hemisphere, ranging from 1.12^oC to 2.22^oC. The differences between global and hemispheric temperatures reflected by the mean estimates of CS are caused by the dissimilarities in land/ocean ratios as suggested in the literature (WGI-IPCC, 2007; Kaufmann and Stern, 2000). But it is worth noting that these differences are not statistically significant, the long run response of a doubling of CO2 concentrations would produce an almost homogeneous warming over both hemispheres.

The estimates presented in Table 14 do not take into account a large part of the feedback processes that exist in the climate system and that are incorporated in complex climate models. Therefore, they could seriously underestimate the CS and the hemispheric differences could be larger due to the diverse regional climate feedback processes. To illustrate the importance that these feedbacks can have over CS estimates, consider the following equation for expressing the climate sensitivity, $S = \Delta T_0/(1-f)$ where f is the feedback factor, which amplify (if $0 < f < 1$) or damp the initial blackbody response of $\Delta T_0 = 1.2°C$ for a CO2 doubling (Knutti and Hegerl, 2008). The typical value for current General Circulation Models has a mean of 0.65 (Knutti and Hegerl, 2008). Using this value for f, climate sensitivity for global temperature would be 4.9^oC instead of 1.7^oC. The underestimates of the CS presented in Table 14 when compared to the CS range produced by climate models are probably due to several feedback processes that a climate model considers but that statistical models exclude. In addition, as argued above, observed temperature records do not provide information about equilibrium states but merely about a climate transient. It is likely that a large part of the warming induced by changes in radiative forcing during the 20th century has not yet manifested itself in observed temperatures (Hansen et al., 1985; IPCC, 2007).

6 Comparison with the cointegration model of Kaufmann et al. (2006a).

Kaufmann et al. (2006a) proposed an alternative model for global temperatures. They assumed the temperature and forcing series to be integrated processes without breaks and accordingly considered an error-correction model based on a cointegrating relation between temperature and forcing labelled CECM. We updated their results using the specification presented in Table III of Kaufmann et al. (2006a). The radiative forcing trend is RFAGG which is composed of the radiative forcing of well-mixed greenhouse gases, the direct and indirect radiative forcing of anthropogenic sulfur emissions and the radiative forcing of solar irradiance. The cointegrating relation between temperatures and RFAGG is given by

$$T_t = \alpha_1 + \beta_1 RFAGG_t + \mu_t \tag{3}$$

so that μ_t is stationary. Efficient estimates of the parameters α_1 and β_1 are obtained using Dynamic Ordinary Least Squares (DOLS) with one lag and lead terms. The results are presented in Table 15. In all cases the long-run relationship between temperatures and radiative forcing represented by β_1 is found to be highly significant and its magnitude is similar across the various temperature series. The estimates $\hat{\alpha}_1$ and $\hat{\beta}_1$ can then be used to estimate the following error-correction model by OLS:

$$\begin{aligned} \Delta T_t &= \alpha_2 + \rho\left(T_{t-1} - \left(\hat{\alpha}_1 + \hat{\beta}_1 RFAGG_{t-1}\right)\right) + \sum_{i=1}^{k} \delta_i \Delta T_{t-i} + \sum_{i=1}^{j} \phi_i \Delta RFAGG_{t-i} \\ &\quad + \sum_{i=1}^{l} \gamma_i SOI_{t-1} + \sum_{i=1}^{m} \psi_i NAO_{t-1} + \mu_1 STRAT\_AER_t + \varepsilon_t \end{aligned} \tag{4}$$

The number of lags were specified as in column (4) of Table III in Kaufmann et al. (2006a), that is $k = 2$, $j = 2$, $l = 1$ and $m = 2$. Note that the CECM in Kaufmann et al. (2006a) was tailored for global temperatures only and not for hemispheric temperatures. However, the results shown below reveal that this specification yields similar results and goodness of fit for NH and SH. Hence, we shall adopt the same model for all temperature series.

The parameter estimates of the ECM (4) estimated by OLS are presented in Table 16. They show that many estimates are not statistically significant. Note that with the exception of the contemporary value of SOI all other series included to account for natural variability are not statistically significant, in particular NAO is not significant in the regression for the NH temperatures and it has a negative sign in the regression for global temperatures. The later results are somewhat at odds with what could be expected from climate theory (Hurrell, 1996). These updated results

are similar to those presented in Kaufmann et al. (2006a). This could be interpreted as a sign of poor model specification. Nevertheless, since our goal is to compare our trend-stationary model, labelled TSM, with the CECM proposed by Kaufmann et al. (2006a), we shall keep the same model specification. Finally, it is interesting to note that the coefficients on the lagged values of $\Delta RFAGG_t$ are not statistically significant in neither Table 16 nor in Table III in Kaufmann et al. (2006a). This gives additional support to our results in Section 4 which stated that the radiative forcing trend does not contribute to the short-term dynamics of temperature and that there are no significant short-term feedback processes, at least when considering an annual frequency.

To compare the two models, we consider both their relative in-sample and out-of-sample forecast performance. The trend stationary models (TSM) considered is the one in Table 13 using $ALL\_FORCING$ and $STRAT\_AER$. We report the root mean squared error (RMSE), the Theil Inequality Coefficient (TU), and to assess whether one model has significantly better predictive power, we use the Diebold-Mariano test (DM; Diebold and Mariano, 1995) and the S_{2a} test (Lehmann, 1975). With $\hat{y}_t$ the forecast of the relevant temperature value y_t, the RMSE is computed as $\sum_{t=t_f}^{t_f+h} (\hat{y}_t - y_t)^2 / h$„ where t_f is the starting date of the forecasting period and h is the forecast horizon. The TU statistic is defined by

$$TU = \left[\frac{(\sum_{t=t_f}^{t_f+h} (\hat{y}_t - y_t)^2 / h)}{(\sum_{t=t_f}^{t_f+h} \hat{y}_t^2/h)(\sum_{t=t_f}^{t_f+h} y_t^2/h)} \right]^{1/2}$$

and has the advantage of being scale invariant and defined between 0 and 1, with zero indicating a perfect forecast. The S_{2a} statistic is defined by:

$$S_{2a} = \frac{\sum_{t=1}^{N} I_+ (d_t) - 0.5N}{\sqrt{0.25N}},$$

where

$$d_t = [y_t - \hat{y}_{t_{TSM}}]^2 - [y_t - \hat{y}_{t_{CECM}}]^2,$$

with $I_+ (d_t) = 1$ if $d_t > 0$, 0 otherwise, N is the number of forecasts and $\hat{y}_{t_{TSM}}$ denotes the forecast from the TSM model while $\hat{y}_{t_{CECM}}$ is the forecast from the CECM model. The DM statistic is the t-statistic of the OLS estimate of α in the following regression:

$$d_t = \alpha + e_t$$

Significantly negative values of the statistics S_{2a} and DM provide evidence that the the TSM forecasts are statistically more accurate than those of the CECM.

Tables 17 and 18 present comparisons of the TSM and CECM forecast performances using in-sample one-step ahead (static) and dynamic forecasts starting from the beginning of the sample to

the last available observation. The results are similar and show that the TSM forecasts for G and NH are more accurate than those of the CECM and that the differences in forecast performances are statistically significant at the 5% level. These results hold for both the HadCRUT3 and the NASA temperature datasets. For SH temperatures, the forecast performances of the TSM and the CECM are similar, and neither the S_{2a} nor the DM tests indicate a statistically significant difference at the 5% level, with the exception of the dynamic forecast using the NASA dataset in which case the CECM performs better.

Table 19 presents the results for an out-of-sample dynamic forecast evaluation. To construct the forecasts, the sample period used for model estimation was 1880-1990 and 1991 to 2010 was the forecasting period with an horizon h of 20 years. The actual values of the explanatory variables were used to generate the forecasts. The results confirm that the forecasts for G and NH from the TSM are significantly more accurate that those from the CECM at the 5% level, and that in the case of SH the differences in forecast performance are not statistically significant.

In summary, the TSM model not only better represents the data generating processes of temperatures but also produces more accurate forecasts for all temperature series with the exception of SH. in which case the forecast performances of TSM and CECM are in general not statistically different. These results are robust to the temperature dataset used and also to the type of forecasts conducted. These findings are consistent with the fact that the TSM model provides a better in sample fit for G and NH temperatures and a nearly equal fit for SH as indicated by the standard errors of the regressions (compare Tables 13 and 16).

7 Conclusions

This paper demonstrates the existence of a robust long-run relationship between radiative forcing and global and temperature series, and that the anthropogenic emissions of greenhouse gases have contributed dominantly to the observed warming during the recorded period. The attribution of climate change to anthropogenic activities presented here is robust in the sense that it takes into account the most important natural and anthropogenic forcing variables, the most important natural variability modes referred in the literature, and is based on the two most important temperature records available.

The univariate time series analysis shows that both temperature and forcing variables are better represented as trend stationary variables with a break in the slope of their trend function and therefore share the same order of integration. These results are congruent with those of Gay et al. (2009) and Estrada et al. (2011) for temperature series, and provide strong evidence for arguing that forcing variables have been erroneously classified as I(1) or I(2) variables in previous attribution

studies based on time series analysis. The time-series properties of all temperature series indicate that during the 20th century their long-run path experienced a structural change in the slope of the trend function that can hardly be interpreted as a part of natural variability. Furthermore, it is shown with multivariate models that these piecewise linear long-term trends cannot be reproduced in a physically acceptable way unless the well-mixed greenhouse gases forcing is included; the other forcing variables by themselves cannot explain the observed warming. The temperature series show a delayed but more than proportional response to the change in the slope of the forcing trend. In one of the cases analyzed, a doubling in the rate of growth of the forcing trend leads to a post-break slope in temperature series of 7.2, 9.1 and 10 times the pre-break slope values for global, northern and southern hemisphere temperatures. The delay in the response was estimated to be 8 years for the southern hemisphere temperature and 24 and 37 years for global and northern hemisphere temperatures. The large delay in the northern hemisphere is due to the low frequency natural variability produced by ocean/atmosphere processes over the North Atlantic (Estrada et al., 2012).

The tests for a common trend demonstrate 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. Some important features that have been described in the literature, such as the cooling of the 1940-1970 period and the slowdown in the warming of the two last decades, are shown to be imprinted in temperatures by the well-mixed greenhouse forcing. Furthermore, the observed warming has clearly been produced by greenhouse gases: about 0.85^oC of the observed warming can be attributed to well-mixed greenhouse forcing, while solar forcing had a warming effect of about 0.06^oC until the first half of the 20th century and then a cooling effect of a similar magnitude. The joint contribution of all forcing variables except well-mixed greenhouse gases produced a net cooling effect of 0.1^oC from 1880 to 2010. It is also shown by means of Granger-causality tests that while forcing variables determine the long-term evolution of global and hemispheric temperatures they do not contribute to the short-term dynamics and that there are no significant short-term feedback processes, at least at the annual frequency.

Parsimonious models for the global and hemispheric temperatures that are able to capture both the short and long-term dynamics are presented. These were shown to provide a better fit and more accurate forecacts than the error correction model of Kaufmann et al. (2006a). The secular movements in global and hemispheric temperatures are represented by the forcing trends that were shown to share their long-term trend with the temperature variables, while the short-run dynamics are influenced by the internal dynamics of temperature series and by the SOI, NAO and AMO which are some climate indices that are known to have an important role determining the inter-annual variability of temperatures. These statistically adequate models provide an excellent

fit to the global and hemispheric temperature series and important similarities and differences in the short-run dynamics and in the secular movement. In all cases, the forcing trend variable is highly significant and its effect on temperatures is larger over the northern hemisphere than in the southern hemisphere or globally. While stratospheric aerosols are shown to have had a large but short-term cooling effect in the northern hemisphere, the effect of volcanic eruptions during the period studied is shown not to have been significant in the southern hemisphere. Short-run dynamics are also shown to be very different in the northern and southern hemispheres. The northern hemisphere is characterized by a much more complex short-run dynamics with stronger persistence and a large influence of ENSO, NAO and AMO. The southern hemisphere temperature has a simpler dynamic in which the excursions from its secular trend are smaller and are mainly influenced by ENSO.

Climate sensitivity estimates are presented for a doubling in the CO2 atmospheric concentrations. These estimates are within the 1.5°C-4.5^{o}C range commonly accepted but as discussed they should be interpreted with care because observed climate is not at an equilibrium state but a climate transient and the statistical models do not include all the relevant feedback processes which are known to have a major role in determining the equilibrium temperature. This paper provides strong evidence on the direct attribution of climate change to anthropogenic activities using observed temperatures, forcing variables and statistical methods.

Appendix

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

Perron (1989) showed that the presence of structural changes can have considerable implications when investigating time-series properties by means of unit root tests. This creates a circular problem given that most of the tests for structural breaks require to correctly identify if the data generating process is stationary or integrated. Depending on whether the process is stationary or integrated the limit distribution of these tests are different and, if the process is misidentified, the tests will have poor properties.

The Perron 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 or contains an autoregressive unit root. 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 model 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 the original Perron-Yabu article. Consider the following data generating process:

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

for $t = 1, ..., T$, $e_t \sim i.i.d.\left(0, \sigma^2\right)$, 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 in the next paragraphs. The initial condition u_0 is assumed to be bounded in probability. The autoregressive coefficient is such that $|\alpha| \leq 1$ and therefore, both integrated and stationary errors are allowed. The interest is testing the null hypothesis $R\Psi = \gamma$ where R is a $(q \times r)$ full rank matrix and 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 $T_1 = [\lambda T]$ for some $\lambda \in (0, 1)$, where $[\cdot]$ denotes the largest integer that is less than or equal to the argument and $1(\cdot)$ is the indicator function.

The model to test for a one-time change in the slope of the trend function is specified with $x_t = (1, t, DT_t)'$ and $\Psi = (\mu_0, \beta_0, \beta_1)'$ where $DT_t = 1(t > T_1)(t - T_1)$ so that the trend function is joined at the time of the break. The hypothesis of interest is $\beta_1 = 0$. The testing procedure is based on a Quasi Feasible Generalized Least Squares approach that uses a superefficient estimate of the sum of the autoregressive parameters 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 $\hat{\alpha}_M$;

4. Apply the truncation
$$\hat{\alpha}_{MS} = \begin{cases} \hat{\alpha}_M & if \quad |\hat{\alpha}_M - 1| > T^{-\frac{1}{2}} \\ 1 & if \quad |\hat{\alpha}_M - 1| \leq T^{-\frac{1}{2}} \end{cases}$$

5. Apply a Generalized Least Squares (GLS) procedure with $\hat{\alpha}_{MS}$ to obtain the estimates of the coefficients of the trend and the variance of the residuals and construct the standard Wald-statistic W_{FMS};

6. Since the break date is assumed to be unknown, the 5 steps above must be repeated for all permissible break dates to construct the Exp functional of the Wald test denoted by
$$Exp\text{-}W_{FS} = \log\left[T^{-1}\sum_{\Lambda}\exp\left(\frac{1}{2}W_{FMS}\left(\lambda\right)\right)\right]$$
where where $\Lambda = \{\lambda; \epsilon \leq \lambda \leq 1 - \epsilon\}$ for some $\epsilon > 0$. We set $\epsilon = 0.15$ as is common the literature.

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

As shown in Perron (1989), the sum of the first order autoregressive coefficients is highly biased towards unity if there is a shift in the trend function. In this case, the unit root null is hardly rejected even if the series is composed of *i.i.d.* disturbances around a trend. Furthermore, if the break occurs in the slope of the trend function, unit root tests are not consistent, i.e., the null hypothesis of a unit root cannot be rejected even asymptotically.

Perron (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 in the "changing growth" model, which can be briefly described as follows. The null hypothesis is:

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

where $DU_t = 1$ if $t > T_B$, 0 otherwise; T_B refers to the time of the break, and $A(L)\, e_t = B(L)\, v_t$, $v_t \sim i.i.d.\left(0, \sigma^2\right)$, with $A(L)$ and $B(L)$ pth and qth 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:

$$y_t = \mu + \beta_1 t + \beta_2 DT_t^* + \widetilde{y}_t \tag{A.2}$$

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

$$\widetilde{y}_t = \alpha \widetilde{y}_{t-1} + \sum_{i=1}^{k} c_i \Delta \widetilde{y}_{t-i} + e_t \tag{A.3}$$

where the k lagged values of $\Delta \widetilde{y}_{t-i}$ are added as a 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 (A.2). The resulting unit root test is then the t-statistic for testing that $\alpha = 1$ in regression (A.3) estimated by OLS. The critical values of the limit distribution of the test are tabulated in Perron (1997).

A problem with most procedures for testing for unit roots in the presence of a one-time break that occurs at an unknown date is that the change in the trend function is allowed only under the alternative hypothesis of a stationary noise component. 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 the 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 $\hat{T}_B = \hat{\lambda}T$ by minimizing the sum of squared residuals using regression (A.2). Then construct a window around that estimate defined by a lower bound T_l and an upper bound T_h. A window of 6 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 to $T_l + 1$ to T_h, and shifting down the data after the window by $S(T) = y_{T_h} - y_{T_l}$; hence,

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

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

$$\widetilde{y}_t^n = \widetilde{\alpha}\widetilde{y}_t^n + \sum_{i=1}^{k} c_i \Delta \widetilde{y}_{t-i}^n + \widetilde{e}_t \tag{A.4}$$

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

The number of lags in (A.3) and (A.4) was chosen using the Schwarz Information Criterion (BIC) but the results are robust to alternative methods for choosing the lag length such as the Akaike Information Criterion (AIC) or the Hannan-Quinn criterion (HQ). After choosing the lag length with the information criteria, Ljung-Box tests were performed on the residuals. In all cases, no evidence of remaining autocorrelation was found.

3 Perron-Zhu methodology for estimating 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. Their results are of particular relevance when considering the problem of selecting the break date when testing for structural changes and for deriving the limiting distributions of unit root tests that allow for a one-time structural change that occurs at an unknown date, such as in the Kim and Perron (2009) test. 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. Our case pertains to the first specification with a stationary noise component. The interested reader is referred to Perron and Zhu (2005) for

the specifications and limiting distributions for the other models considered by the authors. The deterministic part is specified as:

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

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

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

with $T_1 = \lambda T$ the break date and λ the break fraction. 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_1. This specification is therefore referred to as the "joint broken trend". The estimation method is simply to select the break date that minimizes the sum of squared residuals from a regression of the series of interest y_t $(t = 1, ..., T)$ on the regressors $\{1, t, B_t\}$, i.e., applying OLS to the model

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

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

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

where β_b^0 is the true value of the change in the slope parameter and $\sigma^2 = var(u_t)$. Hence, it is straightforward to construct a confidence interval using the estimates of σ^2, β_b^0 and λ_0.

Note that the limiting distribution of the break date does not depend on the structure of the errors (except from the variance term σ^2), in particular it does not depend on the nature of the serial correlation. 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 Testing for a common long-run path

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

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

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

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

If the individual nonstationary components d_t and w_t are present in y_t and x_t but not in ε_t, that is the residuals are found to be stationary, then it is said that y_t and x_t share the same long-run path. The existence of remaining non-common nonstationarities can be evaluated by applying standard unit root tests. This procedure is similar to the Engle and Granger (1987) two-steps cointegration test but it does not require the assumption of unit roots in radiative forcing and temperature variables. Also, since all series are trend-stationary, the relevant critical values are those tabulated for standard unit root (or stationarity) tests with no deterministic terms included.

Since one of the principal issues investigated in this paper is the possibility of radiative forcing and temperature variables sharing common breaks in the slope of the trend function, we also applied a more specific test for evaluating it. Following the same rational discussed above and the concept of common features of Englea and Kozicki (1993), the Perron and Yabu (2009b) test was used for evaluating if breaks that were previously detected in the series y_t and x_t were still present in the residuals ε_t.

The results from the Perron and Yabu test (2009b) and from the standard unit root tests applied to the residuals can provide sound evidence regarding the existence of a common long-term path, as well as of some associated common stylized facts such as breaks and transition periods.

5 Bierens nonparametric nonlinear co-trending test

Nonlinear co-trending is 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 by Bierens (2000) can be expressed as follows:

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

with

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

where z_t is a k-variate time series, u_t is a k-variate zero-mean stationary process and $f(t)$ is deterministic k-variate general nonlinear trend function that 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-trended, implying that there is one or more linear combinations of the time series that are stationary around a constant or a linear trend. Note that this test is a cointegration test in the case when it is applied to series that contain unit roots.

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

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

where

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

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

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

where

$$\hat{F}(x) = \frac{1}{m}\sum_{j=0}^{m-1}\left(z_{[nx]+1-j} - \hat{\beta}_0 - \hat{\beta}_1(nx+1-j)\right)$$

if $[nx] \geq m-1$, $\hat{F}(x) = 0$ if $[nx] < m-1$, $m = n^{\alpha}$ with n equal to 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 solution $\hat{\lambda}_r$, the test statistic is $n^{1-\alpha}\hat{\lambda}_r$. The null hypothesis is that there are r co-trending vectors against the alternative of $r-1$ co-trending vectors. This test has a non-standard distribution and the critical values have been tabulated by Bierens (2000). The existence r co-trending vectors in $r+1$ series indicates the presence of r linear combinations of the series that are stationary around a linear trend and that these series share a single common nonlinear deterministic trend. Such a result indicate a strong secular co-movement in the $r+1$ series.

References

Andrews, D.W.K., and Ploberger, W. (1994) Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative. Econometrica 62, 1383-1414.

Bierens, H. J. (2000) Nonparametric nonlinear cotrending analysis, with an application to interest and inflation in the United States. JBES, 18, 323-337.

Box, G. E. P. & Tiao G. C. (1975) Intervention Analysis with Applications to Economic and Environmental Problems. J. Amer. Statistical Assoc., 70, 70-79.

Brohan P., Kennedy J.J., Harris I., Tett S.F.B., Jones P.D. (2006) Uncertainty estimates in regional and global temperature changes: a new dataset from 1850, Journal of Geophysical Research 111:D12106.

Dickey D.A., Fuller W.A. (1979) Distribution of the estimators for autoregressive time series with a unit root, J. Am. Statist. Assoc. 74, 427–431.

Diebold F. X., Mariano R. S. (1995) Comparing predictive accuracy, JBES 13, 253–263.

Enfield, D.B., Mestas-Nunez A.M., Trimble, P.J. (2001) The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the continental U.S., Geophys. Res. Lett. 28, 2077-2080.

Engle R.F., Granger C.W.J. (1987) Co-integration and error correction: Representation, estimation and testing, Econometrica 55, 251-276.

Engle, R. F. & Kozicki S. (1993) Testing for Common Features. JBES, 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", Clim. Change 101, 407-414, DOI: 10.1007/s10584-010-9928-0.

Estrada F., Perron P., Gay C., Martínez, B. A time series analysis of the 20th century climate simulations produced for the IPCC's AR4. Working paper. Available at http//people.bu.edu/perron/papers/EPGM_2011_web.pdf.

Estrada F., Perron P., Martínez, B. (2012) Statistical evidence about the human impact on climate change. Working paper, Department of Economics, Boston University.

Gay C., Estrada F., Sanchez A. (2009) Global and hemispheric temperature revisited, Clim. Change 94, 333–349, DOI:10.1007/s10584-008-9524-8.

Gil-Alana L.A. (2008) Time trend estimation with breaks in temperature time series, Climatic Change 89:325–337.

Hansen, J.E., Lebedeff, S. (1987) Global trends of measured surface air temperature, J. Geophys. Res. 92, 13345-13372, doi:10.1029/JD092iD11p13345.

NATO ASI Series Vol. I 42. Springer-Verlag. Heidelberg, Germany, 233-272.

Hansen, J., Ruedy, R., Sato, M., Lo, K. (2010), Global surface temperature change, Rev. Geophys. 48, RG4004, doi:10.1029/2010RG000345.

Hansen J., Sato, M., Kharecha, P., von Schuckmann, K. (2011) Earth's energy imbalance and implications, arXiv:1105.1140.

Hansen, J., et al. (1996) A Pinatubo climate modeling investigation. In The Mount Pinatubo Eruption: Effects on the Atmosphere and Climate (G. Fiocco, D. Fua, and G. Visconti, Ed.).

Hasselmann, K. (1979) On the signal-to-noise problem in atmospheric response studies. In: Meteorology of Tropical Oceans [Shaw, D.B. (ed.)]. Royal Meteorological Society, Bracknell, UK, 251–259.

Hasselmann, K. (1997) Multi-pattern fingerprint method for detection and attribution of climate change. Climate Dyn. 13, 601–612.

Houghton, J.T., Jenkins, G.J., Ephraums, J.J. (1990) Climate Change: The IPCC Scientific Assessment, Cambridge University Press, Cambridge, U.K..

Hurrell, J.W. (1995) Decadal trends in the North Atlantic Oscillation and relationships to regional temperature and precipitation, Science 269, 676-679.

Hurrell, J. W. (1996) Influence of variations in extratropical wintertime teleconnections on northern hemisphere temperature, Geophys. Res. Lett. 23, 665–668.

IPCC-WGI: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY (2007).

Jones, P.D., Raper, S.C.B., Bradley, R.S., Diaz, H.F., Kelly, P.M., Wigley, T.M.L. (1986a) Northern hemisphere surface air temperature variations: 1851-1984. Journal of Climate and Applied Meteorology 25,161-179.

Jones, P.D., Raper, S.C.B., Wigley, T.M.L. (1986b). Southern hemisphere surface air temperature variations: 1851-1984. Journal of Climate and Applied Meteorology 25,1213-1230.

Jones, P.D., Wigley, T.M.L., Wright, P.B. (1986c). Global temperature variations between 1861 and 1984, Nature 322, 430-434.

Kaufmann, R.K., Kauppi, H., Stock, J.H. (2006a) Emissions, concentrations, & temperature: A time series analysis, Clim. Change DOI: 10.1007/s10584-006-9062-1.

Kaufmann, R.K., Kauppi, H., Stock, J.H. (2006b) The relationship between radiative forcing and temperature: What do statistical analyses of the instrumental temperature record measure? Clim. Change, 10.1007/s10584-006-9063-0.

Kaufmann, R.K., Stern, D.I. (1997) Evidence for human influence on climate from hemispheric temperature relations, Nature 388, 39–44.

Kerr, R.A. (2000) A North Atlantic climate pacemaker for the centuries, Science 288, 1984–1985.

Kim, D., Perron P. (2009) Unit root tests allowing for a break in the trend function under both the null and the alternative hypotheses, J. Econom. 148, 1-13.

Knudsen, M.F., Seidenkrantz, M.S., Jacobsen, B.H., Kuijpers, A. (2010) Tracking the Atlantic Multidecadal Oscillation through the last 8,000 years, Nature Communications 2:178, DOI: 10.1038.

Lehmann, E. L. (1975) Nonparametrics, statistical methods based on ranks. McGraw-Hill, San Francisco.

McCormick, M.P., Thomason, L.W., Trepte, C.R. (1995) Atmospheric effects of the Mt Pinatubo eruption, Nature, 373.

Mills, T.C. (2010a) 'Skinning a cat': alternative models of representing temperature trends. An editorial comment, Clim. Change 101, 415-426, DOI: 10.1007/s10584-010-9801-1.

Mills T.C., (2010b) Is global warming real? Analysis of structural time series models of global and hemispheric temperatures, Journal of Cosmology 8.

Mitchell, J.F.B., 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, J.T., et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, 695–738.

Perron, P. (1989) The great crash, the oil price shock, and the unit root hypothesis, Econometrica 57, 1361–1401.

Perron, P., (1997) Further evidence on breaking trend functions in macroeconomic variables, J. Econom 80, 355-385.

Perron, P. (2006) Dealing with structural breaks. In T. C. Mills and K. Patterson (eds.), Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory. Palgrave Macmillan, New York, 278-352.

Perron, P., Vogelsang, T.J. (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 of Stationary Noise Component. J. Econom, 151, 56-69.

Perron, P., Yabu, T. (2009b) Testing for shifts in trend with an integrated or stationary noise component, JBES 27, 369-396.

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

Rosenzweig, C., Karoly, D., Vicarelli, M., Neofotis, P., Wu, Q., Casassa, G., Menzel A., Root, T.L., 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, doi:10.1038/nature06937.

Roy, A., & Fuller, W.A. (2001) Estimation for Autoregressive Processes With a Root Near One. JBES, 19, 482-493.

Said, E., Dickey, D.A. (1984) Testing for unit roots in autoregressive moving average models of unknown order, Biometrika 71, 599–607.

Schlesinger, M.E., Ramankutty, N. (1994) An oscillation in the global climate system of period 65–70 years, Nature 367, 723–726.

Stern, D.I., Kaufmann, R.K. (1997a) Is there a global warming signal in hemispheric temperature series? Working Papers in Ecological Economics, The Australian National University, Center for Resource and Environmental Studies Ecological Economics Programme (available at http://www.bu.edu/cees/research/workingp/pdfs/9903.pdf).

Stern, D.I. Kaufmann, R.K. (1997b) Time series properties of global climate variables: Detection and attribution of climate change, Working Paper in Ecological Economics, Australian National University, Center for Resource and Environmental Studies, Ecological Economics Programme (http://eprints.anu.edu.au/archive/00000665/00/eep9702.pdf).

Stock, J. H. and Watson, M. W. (1993) A simple estimator of cointegrating vectors in higher order integrated systems, Econometrica 61, 783–820.

Stott, P.A., et al. (2006) Transient climate simulations with the HadGEM1 model: causes of past warming and future climate change, J. Clim. 19, 2763–2782.

Tol, R.S.J., De Vos, A.F. (1993) Greenhouse statistics - time series analysis, Theor. Appl. Climatol. 48, 63–74.

Tol, R.S.J., De Vos, A.F. (1998) A Bayesian statistical analysis of the enhanced greenhouse effect, Clim. Change 38, 87–112.

Triacca, U. (2001) On the use of Granger causality to investigate the human influence on climate, Theor. Appl. Climatol. 69, 137–138.

WGII-IPCC, 2007. Climate Change 2007: Impact, Adaptation and Vulnerability, Contribution of the Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson (eds)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY.

Wigley, T.M.L., Raper, S.C.B. (1990) Natural variability of the climate system and detection of the greenhouse effect, Nature 344, 324-327.

Wolter, K., Timlin, M.S. (1998) Measuring the strength of ENSO - how does 1997/98 rank? Weather 53, 315-324.

Zivot, E., Andrews, D.W.K. (2002) Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis, JBES 20, 25-44.

Zwiers F., Hegerl, G. (2008) Climate change: Attributing cause and effect, Nature 453, 296-297, doi:10.1038/453296a.

Table 1. Standard Augmented Dickey-Fuller test applied to temperature series

Series	ADF test statistic
Global (NASA)	-1.49
Global (HadCRUT3)	-2.30
NH (NASA)	-1.75
NH (HadCRUT3)	-2.40
SH (NASA)	-2.92
SH (HadCRUT3)	-2.80

Note: The AIC information criterion was used for determining the lag length, and the deterministic regressors were a constant and a trend. None of the statistics are significant at the 5% level (critical value: -3.41).

Table2. Tests for structural changes in the trend function for time series with an integrated or stationary noise component

Series	Exp-Wald statistic value
Global (NASA)	5.20[a]
Global (HadCRUT3)	4.36[a]
NH (NASA)	4.04[a]
NH (HadCRUT3)	4.93[a]
SH (NASA)	5.14[a]
SH (HadCRUT3)	5.77[a]

Note: a, b, c denote statistical significance at the 1%, 5% and 10%, respectively.

Table 3. Tests for a unit root with a one-time break in the trend function, HadCRUT3 (C) and NASA (N) databases.

Series	T_b	k	$\hat{\mu}$	$t_{\hat{\mu}}$	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\gamma}$	$t_{\hat{\gamma}}$	$\hat{\alpha}$	$t_{\hat{\alpha}}$	$S(\hat{e})$	$t_{\alpha}(\hat{\lambda}_{tr}^{AO})$
G_C	1971	0	-0.40	-18.09	0.0022	7.45	0.0136	10.17	0.59	-6.37[a]	0.126	-5.61[A]
NH_C	1984	0	-0.40	-15.37	0.0030	9.27	0.0243	9.45	0.58	-6.46[a]	0.154	-5.91[A]
SH_C	1955	0	-0.38	-17.03	0.0011	3.12	0.0099	10.26	0.55	-6.86[a]	0.122	-6.26[A]
G_N	1978	0	-0.32	-16.58	0.0039	12.15	0.0129	9.04	0.40	-7.45a	0.091	-6.40A
NH_N	1985	0	-0.31	-11.81	0.0042	10.02	0.0235	9.00	0.46	-6.90[a]	0.124	-5.57[A]
SH_N	1923	0	-0.24	-9.846	-0.0005	-0.70	0.0078	7.85	0.35	-7.91[a]	0.084	-6.82[A]
SH_N*	1955	0	-0.30	-15.40	0.0027	6.84	0.0066	7.56	0.36	-7.79[a]	0.091	-5.91[A]

Note: The regression model for the unit root tests is defined in equations (15) and (16) of the Methods section. The symbols used are defined as follows: Tb is the estimated time of the break; k is the number of lagged differences added to correct for serial correlation; $S(\hat{e})$ is the standard error of the regression; $\hat{\mu}$, $\hat{\beta}$, $\hat{\gamma}$ are the regression coefficients of the trend function and $t_{\hat{\mu}}$, $t_{\hat{\beta}}$, $t_{\hat{\gamma}}$ the corresponding t-statistic values. $\hat{\alpha}$ is the first order autoregressive coefficient and $t_{\hat{\alpha}}$ is the Perron (1997) unit root test statistic. a, b, c, d denote statistical significance at the 1%, 2.5%, 5% and 10%, respectively (critical values from Perron, 1997). $t_{\alpha}\left(\hat{\lambda}_{tr}^{AO}\right)$ is the Kim and Perron (2009) unit root test. A, B, C, D denotes statistical significance at the 1%, 2.5%, 5% and 10% , respectively (critical values from Perron and Vogelsang, 1993, Table 1).

Table 4. Break date estimates and 95% confidence intervals for global and hemispheric temperatures

	NASA	HadCRUT3
Globe	1978 (1971, 1985)	1971 (1962, 1980)
NH	1985 (1979, 1991)	1984 (1977, 1991)
SH	1923 (1914, 1932)	1955 (1945, 1965)
SH*	1955 (1945, 1965)	--

Table 5. Eigenvectors of the first two principal components of the set of radiative forcing series.

Variable	PC1	PC2
WM_GHG	0.340	-0.036
SOLAR	0.272	0.120
STRAT_AER	0.011	0.988
H2O	0.345	-0.041
BC	0.339	-0.015
R_AER	-0.343	0.027
AIE	-0.345	-0.018
O3	0.339	-0.060
S_ALBEDO	0.340	-0.016
LAND_USE	-0.331	-0.049

Table 6. Standard Augmented Dickey-Fuller test applied to the various forcing trend

Series	ADF test statistic
WM_GHG	-2.74
GHG_SOLAR	-2.76
ALL_FORCING	-2.73
STRAT_AER	-4.92
PC1	-2.31
PC2	-4.80

Note: The AIC information criterion was used for determining the lag length, and the deterministic regressors were a constant and a trend, except for PC2 and STRAT_AER for which only a constant was included. Figures in bold denote statistical significance at 5% level.

Table 7. Test for a structural change in the trend function allowing an integrated or stationary noise component.

Series	Exp-Wald statistic value
WM_GHG	62.79[a]
GHG_SOLAR	5.41[a]
ALL_FORCING	5.63[a]
PC1	2.65[b]

Note: a, b, c denote statistical significance at the 1%, 5% and 10%, respectively.

Table 8. Unit root tests applied to the "forcing trend" variables allowing for a one-time break in the trend function

Series	T_b	k	$\hat{\mu}$	$t_{\hat{\mu}}$	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\gamma}$	$t_{\hat{\gamma}}$	$\hat{\alpha}$	$t_{\hat{\alpha}}$	$S(\hat{e})$	$t_{\alpha}(\hat{\lambda}_{tr}^{AO})$
WM_GHG	1960	7	-0.0343	-4.08	0.0105	64.95	0.0351	87.76	0.9009	-4.24b	0.040	-3.96B
GHG_SOLAR	1960	2	-0.1459	-8.20	0.0144	44.50	0.0307	42.37	0.8359	-5.59a	0.022	-4.86A
ALL_FORCING	1960	2	-0.0852	-5.34	0.0064	20.82	0.0221	28.98	0.8365	-4.58a	0.076	-4.24B
PC1	1947	2	-4.0457	-110.44	0.0496	61.45	0.0533	35.52	0.6901	-7.37a	0.162	-5.85A

Note: The regression model for the unit root tests is defined in equations (1), (2) and (3). The symbols used are defined as follows: Tb is the estimated time of the break; k is the number of lagged differences added to correct for serial correlation; $S(\hat{e})$ is the standard error of the regression; $\hat{\mu}$, $\hat{\beta}$, $\hat{\gamma}$ are the regression coefficients of the trend function and $t_{\hat{\mu}}$, $t_{\hat{\beta}}$, $t_{\hat{\gamma}}$ the corresponding t-statistic values. $\hat{\alpha}$ is the sum of the first order autoregressive coefficients and $t_{\hat{\alpha}}$ is the Perron (1997) unit root test statistic. a, b, c denote statistical significance at the 1%, 5% and 10%, respectively (critical values from Perron, 1997). $t_{\alpha}\left(\hat{\lambda}_{tr}^{AO}\right)$ is the Kim and Perron (2009) unit root test. A, B, C denotes statistical significance at the 1%, 5% and 10%, respectively (critical values from Perron and Vogelsang, 1993, Table 1).

Table 9. Break date estimates and 95% confidence intervals for the forcing trend variables.

	Break date	
WM_GHG	1960	(1958, 1962)
GHG_SOLAR	1960	(1957, 1963)
ALL_FORCING	1960	(1955, 1965)
PC1	1947	(1944, 1950)

Table 10. Estimated slope coefficient, goodness of fit measure, and ADF test on the residuals of regression (1) for different forcing trends.

		Global Temperature			NH Temperature			SH Temperature		
	Forcing Trends	β	R^2	ADF	β	R^2	ADF	β	R^2	ADF
NASA DATASET	WM_GHG	0.258	0.81	-3.96	0.283	0.70	-4.05	0.232	0.84	-7.45
		(23.59)		(2)	(17.27)		(1)	(26.01)		(0)
	GHG_SOLAR	0.241	0.82	-4.03	0.264	0.70	-4.05	0.217	0.85	-7.98
		(24.07)		(2)	(17.46)		(1)	(26.74)		(0)
	ALL_FORCING	0.409	0.80	-3.76	0.447	0.68	-3.86	0.371	0.84	-7.72
		(22.86)		(2)	(16.67)		(1)	(25.99)		(0)
	PC1	0.079	0.79	-3.52	0.086	0.67	-3.70	0.071	0.83	-7.47
		(22.14)		(1)	(16.24)		(1)	(25.31)		(0)
	All Forcing Except GHG	-0.664	0.79	-6.07	-0.735	0.69	-4.28	-0.593	0.80	-6.79
		(-22.06)		(0)	(-16.96)		(1)	(-22.85)		(0)
		Global Temperature			N H Temperature			S N Temperature		
	Forcing Trends	β	R^2	ADF	β	R^2	ADF	β	R^2	ADF
HadCRUT3 DATASET	WM_GHG	0.255	0.77	-5.85	0.280	0.71	-5.20	0.231	0.74	-5.97
		(21.05)		(0)	(17.59)		(0)	(19.40)		(1)
	GHG_SOLAR	0.238	0.78	-5.87	0.261	0.71	-5.23	0.215	0.74	-5.95
		(21.14)		(0)	(17.73)		(0)	(19.36)		(1)
	ALL_FORCING	0.409	0.78	-5.82	0.444	0.70	-5.06	0.374	0.76	-6.20
		(21.15)		(0)	(17.18)		(0)	(20.27)		(1)
	PC1	0.077	0.74	-5.28	0.085	0.68	-4.86	0.069	0.70	-4.07
		(18.92)		(0)	(16.39)		(0)	(17.34)		(2)
	All Forcing Except GHG	-0.650	0.74	-5.50	-0.722	0.69	-5.19	-0.579	0.69	-5.77
		(-18.98)		(0)	(-16.92)		(0)	(-1.75)		(0)

Note: Figures in brackets show the t-statistic values while figures in parentheses the number of lags to correct for serial correlation chosen using the BIC criterion (???). The critical values for the ADF test are: -2.583 (1%), -1.943 (5%), -1.615 (10%).

Table 11. Comparisons of the post- and pre-break values of the slope of the trend for the forcing and temperature series for the HadCRUT3 (C) and NASA (N) datasets.

Forcing trend	Post-break slope compared to its pre-break value
WM_GHG	334%
GHG_SOLAR	213%
ALL_FORCING	345%
PC1	107%

Temperature	Post-break slope compared to its pre-break value
G_C	718%
NH_C	910%
SH_C	1000%
G_N	431%
NH_N	660%
SH_N	344%

Table 12. Results of the Granger Causality tests applied to the detrended northern hemisphere, southern hemisphere and global temperatures, and detrended PC1; NASA dataset.

Dependent Variable	Excluded	Test	P-value
NH*	SH*	1.02	0.80
	PC1*	4.34	0.23
	SOI	18.51	0.0003
SH*	NH*	0.88	0.83
	PC1*	2.83	0.42
	SOI	18.84	0.0003
PC1*	NH*	1.02	0.80
	SH*	0.89	0.83
	SOI	4.26	0.24
G*	PC1*	3.90	0.27
	SOI	25.97	0.0000
PC1*	G*	0.08	0.995
	SOI	4.39	0.22

Note: SH and NH represent the southern and northern hemisphere temperatures, respectively. * denotes that the variable has been detrended. G represent global temperatures.

Table 13. Estimated coefficients and coefficient of determination of regression (2), HadCRUT3(C) and NASA (N) databases.

		α	β_1	β_2	ϕ_1	ϕ_2	γ_1	δ_1	η_1	η_2	R^2	$S(\hat{e})$
Pc1 and Pc2 used	G_N	-0.009	0.054	0.019	0.275	--	-0.030	--	0.346	--	0.930	0.0690
		(-1.43)	(10.17)	(2.95)	(4.19)		(5.17)		(8.08)			
	NH_N	0.008	0.038	0.038	0.358	0.196	-0.020	0.060	0.642	-0.325	0.915	0.0889
		(0.98)	(5.64)	(4.59)	(4.69)	(3.16)	(-2.53)	(3.20)	(9.07)	(-3.93)		
	SH_N	-0.030	0.040	0.013	0.445	--	-0.035	--	--	---	0.886	0.0778
		(-3.77)	(6.90)	(1.91)	(6.02)		(-5.45)					
	G_C	-0.054	0.035	--	0.525	--	-0.029	--	0.437	-0.225	0.921	0.0744
		(-5.28)	(6.51)		(7.69)		(-4.55)		(8.37)	(-3.66)		
	NH_C	-0.026	0.038	0.021	0.373	0.155	-0.029	--	0.575	-0.253	0.917	0.0885
		(2.70)	(5.11)	(2.51)	(4.41)	(2.41)	(-3.78)		(8.64)	(-3.03)		
	SH_C	-0.081	0.031	--	0.535	--	-0.041	--	0.152	--	0.863	0.0903
		(-6.06)	(6.31)		(8.66)		(-5.64)		(3.36)			
All Forcing and START_ARE Used	G_N	-0.147	0.283	0.053	0.262	--	-0.028	--	0.327	--	0.933	0.0678
		(-8.59)	(10.47)	(4.30)	(4.06)		(-4.88)		(7.98)			
	NH_N	-0.101	0.264	0.096	0.203	0.158	-0.029	0.040	0.479	--	0.910	0.0915
		(-5.49)	(8.60)	(5.64)	(2.85)	(2.56)	(-3.70)	(2.13)	(7.99)			
	SH_N	-0.134	0.210	0.033	0.426	--	-0.033	--	--	--	0.887	0.0775
		(-6.13)	(6.92)	(2.53)	(5.64)		(-5.14)					
	G_C	-0.169	0.212	0.033	0.448	--	-0.031	--	0.407	-0.186	0.930	0.0705
		(-6.60)	(7.16)	(2.53)	(6.11)		(-5.02)		(7.98)	(-3.12)		
	NH_C	-0.123	0.203	0.059	0.370	0.154	-0.026	0.041	0.602	-0.282	0.912	0.0857
		(-4.41)	(5.37)	(3.66)	(4.48)	(2.44)	(-3.28)	(2.27)	(8.76)	(-3.26)		
	SH_C	-0.204	0.197	--	0.443	--	-0.039	--	0.159	--	0.872	0.0872
		(-7.64)	(7.18)		(6.63)		(-5.56)		(3.62)			

Table 14. Climate sensitivity estimates for a doubling of CO2 atmospheric concentrations; HadCRUT3 (C) and NASA (N) datasets.

Dataset	5th percentile	Mean	95th percentile
G_C	1.21	1.72	2.32
G_N	1.35	1.69	2.08
NH_C	1.17	1.98	3.14
NH_N	1.32	1.85	2.52
SH_C	1.12	1.58	2.10
SH_N	1.16	1.63	2.22

Table 15. Estimates of the cointegration relation between temperature series and RFAGG using the DOLS estimator with one lag and lead terms.

	α_1	β_1
G_N	-0.2679**	0.4350**
NH_N	-0.2553**	0.4799**
SH_N	-0.2804**	0.3900**
G_C	-0.3704**	0.4373**
NH_C	-0.3357**	0.4763**
SH_C	-0.4050**	0.3985**

Note: **,* denote statistical significance at 5% and 10% levels, respectively. The standard errors were estimated using the Newey-West heteroskedasticity-robust procedure.

Table 16. Estimates of the error correction model using the specification of Kaufmann et al. (2006a).

	α_2	ρ	δ_1	δ_2	φ_1	φ_2	γ_0	γ_1	ψ_0	ψ_1	ψ_2	μ_1	$S(\hat{e})$
G_N	0.025**	-0.42**	-0.22**	-0.12*	-0.27	0.08	-0.036**	-0.014	-0.014	0.001	-0.016	0.075**	0.081
NH_N	0.036**	-0.34**	-0.28**	-0.07	-0.24	-0.06	-0.039**	-0.023**	0.006	0.023	-0.021	0.124**	0.111
SH_N	0.013	-0.50**	-0.15*	-0.13	-0.23	0.24	-0.033**	-0.008	-0.031**	-0.021	-0.012	0.026	0.077
G_C	0.017*	-0.34**	-0.15	-0.16**	-0.23	0.12	-0.046**	0.002	-0.023	0.005	-0.023	0.055**	0.085
NH_C	0.029**	-0.29**	-0.18*	-0.10	-0.10	-0.26	-0.047**	-0.005	-0.006	0.019	-0.027	0.086**	0.109
SH_C	0.006	-0.38**	-0.15	-0.24**	-0.32	0.52*	-0.045**	-0.003	-0.037**	-0.005	-0.022	0.026	0.087

Table 17. In-sample static (1-step) forecast comparisons of TSM and CECM.

	$RMSE_{TSM}$	$RMSE_{CECM}$	TU_{TSM}	TU_{CECM}	S_{2a}	DM
G_N	**0.0662**	0.0766	**0.1319**	0.1551	-1.5085*	-2.5506**
NH_N	**0.0886**	0.1052	**0.1538**	0.1836	-2.3959**	-3.1437**
SH_N	0.0760	**0.0732**	0.1694	**0.1612**	0.6212	1.5453
G_C	**0.0686**	0.0817	**0.1224**	0.1465	-3.2832**	-3.3305**
NH_C	**0.0826**	0.1040	**0.1411**	0.1791	-2.3958**	-3.9456**
SH_C	0.0856	**0.0825**	0.1489	**0.1407**	0.0887	0.8271

Note: **,* denote statistical significance at 5% and 10% levels, respectively. Bold figures denote the smallest forecast error statistic.

Table 18. In-sample dynamic forecast comparisons of TSM and CECM.

	$RMSE_{TSM}$	$RMSE_{CECM}$	TU_{TSM}	TU_{CECM}	S_{2a}	DM
G_N	0.0689	0.0827	0.1373	0.1649	-1.6850*	-2.6273**
NH_N	0.0917	0.1212	0.1592	0.2098	-2.0492**	-3.7630**
SH_N	0.0834	0.0781	0.1860	0.1695	1.6860*	2.1102**
G_C	0.0784	0.0985	0.1406	0.1751	-2.2184**	-3.7735**
NH_C	0.0927	0.1345	0.1589	0.2323	-3.2832**	-4.8641**
SH_C	0.1007	0.0953	0.1760	0.1610	1.6860*	0.8895

Note: **,* denote statistical significance at 5% and 10% levels, respectively. Boldfigures denote the smallest forecast error statistic.

Table 19. Out-of-sample dynamic (20-steps) forecast comparisons of TSM and CECM.

	$RMSE_{TSM}$	$RMSE_{CECM}$	TU_{TSM}	TU_{CECM}	S_{2a}	DM
G_N	0.0765	0.1144	0.0846	0.1292	-2.5236**	-2.1039**
NH_N	0.0984	0.1525	0.0931	0.1427	-2.0647**	-2.1600**
SH_N	0.1226	0.1179	0.1730	0.1660	0.2294	1.1881
G_C	0.0709	0.1057	0.0999	0.1510	-1.1471	-2.3514**
NH_C	0.0815	0.1381	0.0897	0.1543	-1.6059*	-3.1927**
SH_C	0.1070	0.1045	0.2039	0.2031	0.2294	0.5913

Note: **,* denote statistical significance at 5% and 10% levels, respectively. Bold figures denote the smallest forecast error statistic.

Figure 1. Time series plot of the global and hemispheric temperatures series from the HadCRUT3 and NASA databases.

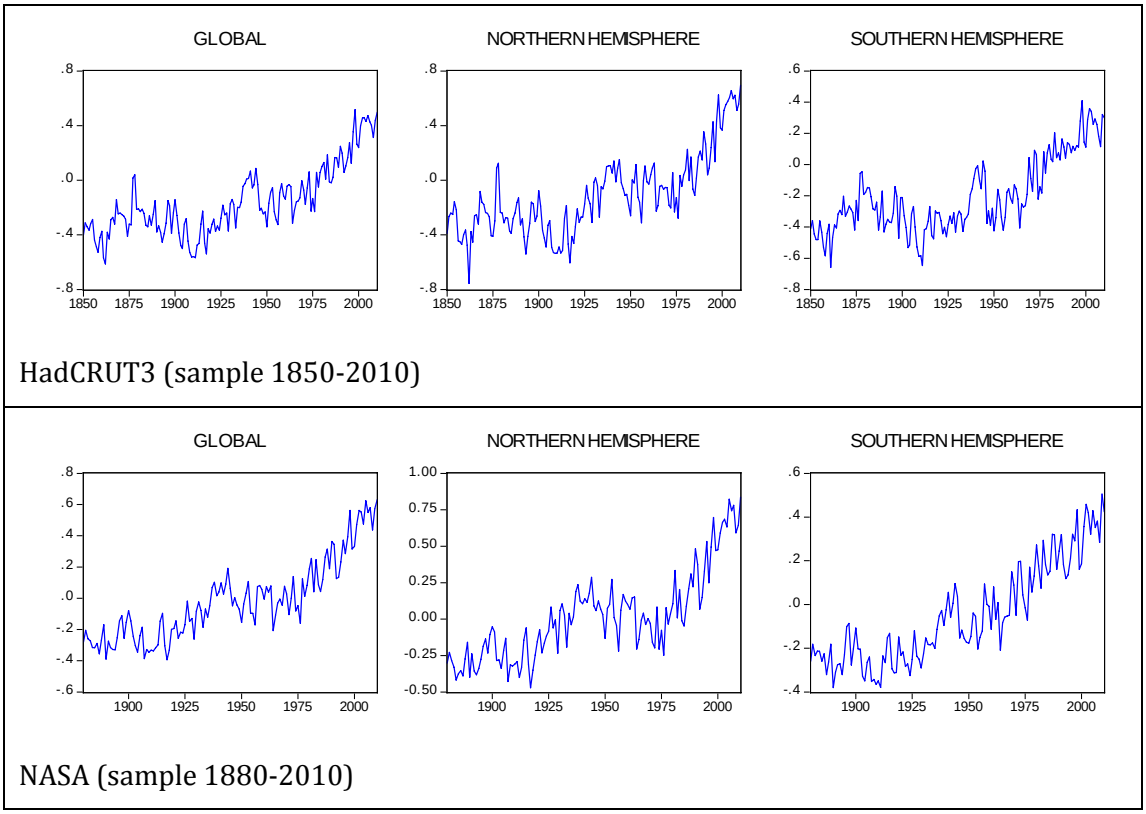

Figure 2. Sum of the squared residuals of the regressions used for estimating the beak date, for the Southern Hemisphere temperature series from the NASA and HadCRUT3 datasets with different samples.

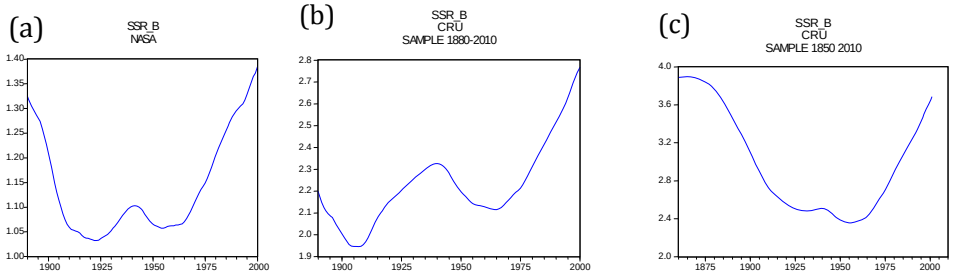

Figure 3. Time series plot of the forcing variables: well mixed greenhouse gases (WM_GHG); ozone (O3); stratospheric H2O (H2O); solar irradiance (SOLAR); land use change (LAND_USE); snow albedo (S_ALBEDO); stratospheric aerosols (STRAT_AER); black carbon (BC); reflective aerosols (R_AER); aerosols indirect effect (AIE). Units are watts per squared meter.

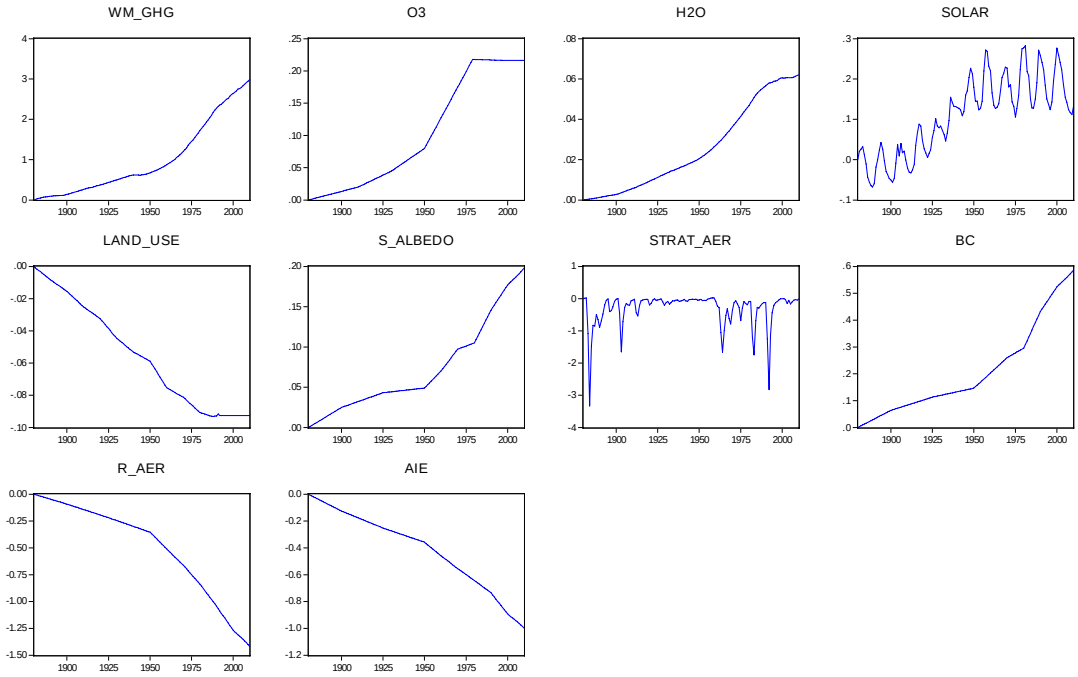

Figure 4. Forcing trends constructed as the sum of the radiative forcing variables; units are watts per squared meter.

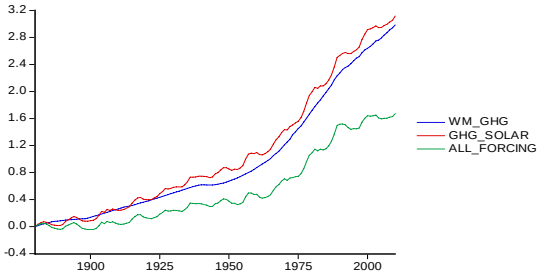

Figure 5. First and second principal components of the set of radiative forcing series.

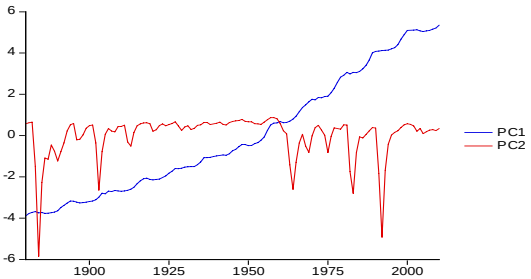

Figure 6. Observed and fitted temperatures using different definitions of the forcing trend

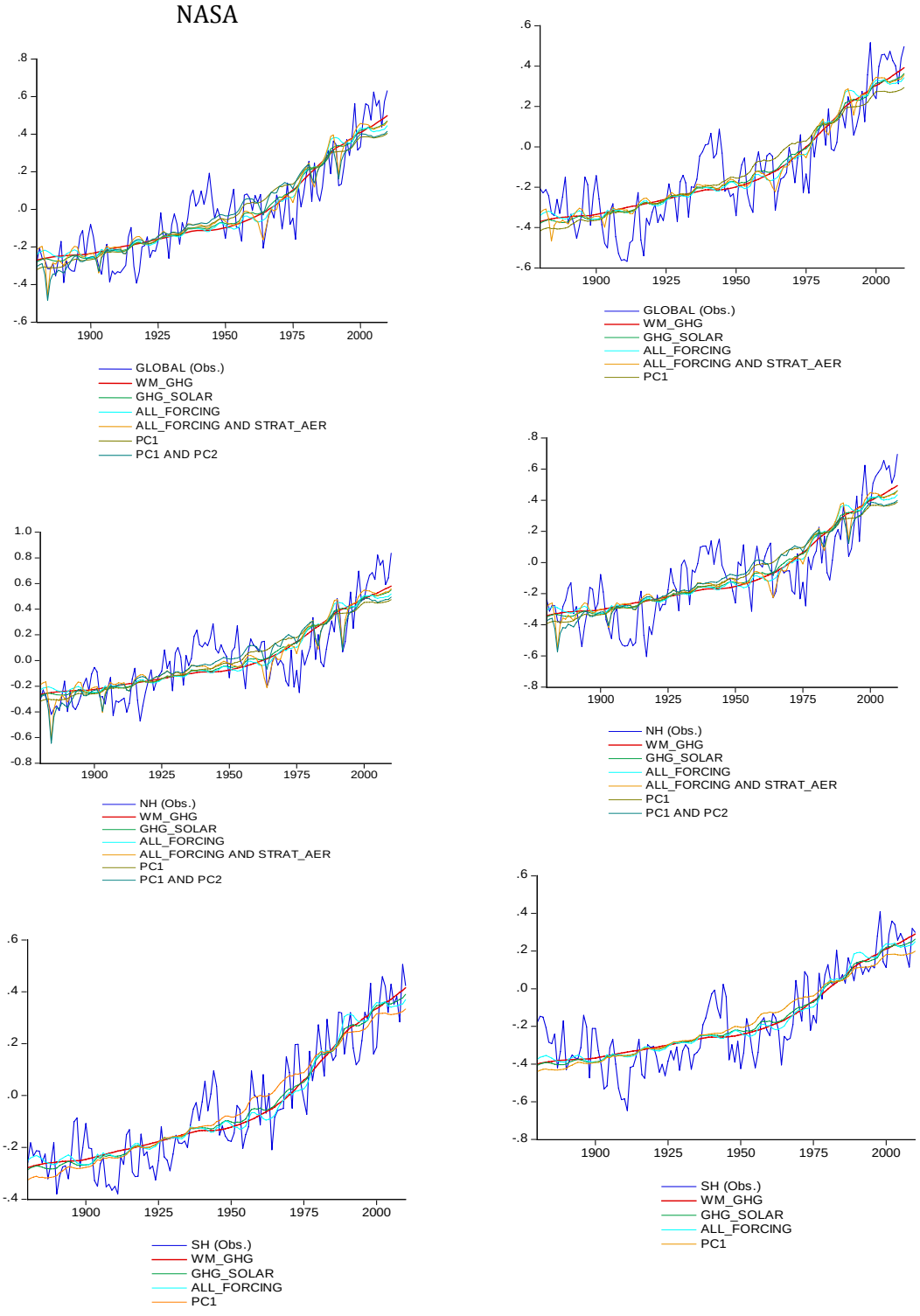

Figure 7. Contributions of different combinations of the forcing trend to the observed warming in the northern hemisphere temperature.