Climatic Change (2009) 94:333–349
DOI 10.1007/s10584-008-9524-8

Global and hemispheric temperatures revisited

Carlos Gay-Garcia · Francisco Estrada · Armando Sánchez

Received: 17 October 2006 / Accepted: 22 September 2008 / Published online: 5 December 2008

Abstract To characterize observed global and hemispheric temperatures, previous studies have proposed two types of data-generating processes, namely, random walk and trend-stationary, offering contrasting views regarding how the climate system works. Here we present an analysis of the time series properties of global and hemispheric temperatures using modern econometric techniques. Results show that: The 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 affected the mean of the processes but not their variability; it has manifested in two stages in global and Northern Hemisphere temperatures during the last century, while a second stage is yet possible in the Southern Hemisphere; in terms of Article 2 of the Framework Convention on Climate Change it can be argued that significant (dangerous) anthropogenic interference with the climate system has already occurred.

C. Gay-Garcia (✉) · F. Estrada
Centro de Ciencias de la Atmósfera, UNAM,
Ciudad Universitaria, Circuito Exterior,
México, DF 04510, México
e-mail: cgay@servidor.unam.mx

F. Estrada
e-mail: feporrua@atmosfera.unam.mx

A. Sánchez
Instituto de Investigaciones Económicas,
Universidad Nacional Autónoma de México,
México, DF 04510, México

1 Introduction

Although there is a growing consensus in the scientific community that the increase in GHG concentrations will lead to changes in temperature and precipitation at regional and global scales, there is still a discussion concerning the detection, attribution and magnitude of climate change in observed global and hemispheric temperatures (IPCC 2001; Folland et al. 1992; Hansen and Lebedeff 1987; Jones and Moberg 2003; Mann et al. 1998, 1999; Jones and Mann 2004; Mann and Jones 2003; Stern and Kaufmann 1997, 1999). This discussion has shown to be of prime interest not only in the scientific field, but also for risk management and policy making (IPCC 2001; Joachim et al. 2006; Bruckner and Schellnhuber 1999; Schneider 2002).

The literature on the stochastic behavior of temperature series has gone through different stages. Initially, it translated into finding statistical evidence in favor of a deterministic trend, against a stochastic trend generated by a unit root process. The first case implies that the recent warming could be interpreted as a long term trend and that global and hemispheric temperatures can be modeled as stationary processes around a deterministic trend that is assumed to be produced by the increase in GHG atmospheric concentrations (Galbraith and Green 1992; Bloomfield 1992; Zheng and Basher 1999). In the latter, the upward trend of global temperature in the last decades should not be predicted to continue and was interpreted as evidence against anthropogenic climate change (Woodward and Gray 1993, 1995; Kärner 1996).

The lack of consensus regarding the adequate representation of non-stationarities present in global and hemispheric temperature series also obscured the possibility of finding a statistically significant relation between global temperature and atmospheric concentrations of GHG gases (Folland et al. 1992). Stern and Kaufmann (1997) provided a breakthrough on this issue by applying cointegration techniques to global and hemispheric temperature series and radiative forcing variables. Cointegration provided a statistically rigorous and direct attribution of climate change to anthropogenic activity, changing the idea of how a stochastic trend in temperature was interpreted (Stern and Kaufmann 1997, 1999; Kaufmann and Stern 1997; Kaufmann et al. 2006a, b). However, here we show that the assumption of a unit root in observed temperature series has strong implications that may not be consistent with the physical processes that govern global temperature. On the other hand, a trend-stationary representation is statistically and physically adequate and provides important inferences regarding how climate change has manifested in global and hemispheric temperatures.

2 Interpretation of unit root and trend stationary processes in temperature series

Following Nelson and Plosser (1982), we will define a trend-stationary (TS) process as a stochastic process that consists of a deterministic trend (any function of time) plus a stationary stochastic process which is assumed to have a representation as a stationary and invertible ARMA process. A trend stationary model can be expressed as:

$$y_t = \alpha + \beta t + c_t, \tag{1}$$

$$\phi(L)\, c_t = \theta(L)\, u_t \tag{2}$$

where $u_t \sim i.i.d\left(0, \sigma_u^2\right)$,[1] α and β are fixed parameters, L is the lag operator, and $\phi(L)$ and $\theta(L)$ are polynomials in L that satisfy the conditions for stationarity and invertibility. The long term or secular growth of this process is dominated by its deterministic part $(\alpha + \beta t)$, while fluctuations are transitory and dissipate over time. Neither present nor past fluctuations will alter the long-run path of the series and the process tends to revert to its conditional mean. This process implies that uncertainty is bounded because it has constant variance and its long-term forecast error has a finite variance (Maddala and Kim 1998; Enders 2003; Hatanaka 1996). In the context of climate modeling, the secular movement or long-term variations (climate changes) are determined by changes in the external forcing factors such as orbital geometry, solar irradiance, greenhouse gas concentrations and changes in other radiative forcing variables. The stochastic stationary component is interpreted as the (shorter term) natural variability and has limited memory. Therefore, the long term forecast is dominated by external forcing factors—as occurs in most physical based climate models, while the process' autocorrelation (persistence) provides important information regarding shorter term forecasting.

The second class of non-stationary processes considered here is referred to as difference-stationary (DS). This is an integrated stochastic process that is non-stationary in levels but for which its first or higher order differences is a stationary and invertible ARMA process. Consider the following first order difference-stationary, unit root plus drift model:

$$y_t = \beta + y_{t-1} + d_t \tag{3}$$

$$\delta\left(L\right) d_t = \lambda\left(L\right) u_t \tag{4}$$

where $u_t \sim i.i.d\left(0, \sigma_u^2\right)$, and $\delta(L)$ and $\lambda(L)$ are polynomials that satisfy the stationarity and invertibility conditions. The basic difference between TS and DS is that the latter also exhibits a secular movement but it is stochastic, not deterministic (Maddala and Kim 1998; Enders 2003; Hatanaka 1996). The stochastic nature of this class of process is clearly illustrated assuming an initial condition and solving iteratively the difference equation:

$$y_t = y_0 + \beta t + \sum_{i=0}^{t-1} d_{t-i} \tag{5}$$

where y_0 is the initial condition, βt is a deterministic trend and $\sum\limits_{i=0}^{t-1} d_{t-i} = v_t$ has a stochastic trend—a hardly predictable but systematic variation (Maddala and Kim 1998)—consisting in the sum of the stationary error term. In this case, the variance of the process[2] $Var\left(y_t\right) = t\sigma_d^2$ increases with time and diverges as $t \to \infty$ (Hatanaka 1996). Two of the most important features of this type of process are that it has no tendency to return to a long-run mean and that it has infinite memory; shocks do not fade (Maddala and Kim 1998; Hatanaka 1996). The sum of these random shocks determines the secular movement of the series. That is, each and every shock

[1] i.i.d. means independent and identically distributed.

[2] If errors are white noise.

produced by natural variability or changes in external forcing factors have permanent effects on the long-run path of temperature. Assuming a unit root implies that global and hemispheric temperatures are highly unstable processes and, therefore, single events such as isolated solar flares, the 1974 La Niña (as well as other internal variations) or the 1883 Krakatau eruption would have changed the long-run path of global temperatures and their effect would be present even today. Moreover, past and present events are in equal footing for determining current conditions: any shock that occurred even in the distant past is as important as present variations. The long-term forecast is always influenced by historical events, the forecast error is not bounded, and it increases as the period to be forecasted is extended. Temperature predictability would be very limited, even if forcing factors are held constant (Gay et al. 2007).

3 Unit root testing and structural change

Econometric literature is vast on the issue of discriminating between trend-stationary and difference-stationary processes given the important implications each of them has for macroeconomic theory. Therefore, over the last 20 to 30 years there has been much progress in testing for a unit root (Stock 1994).

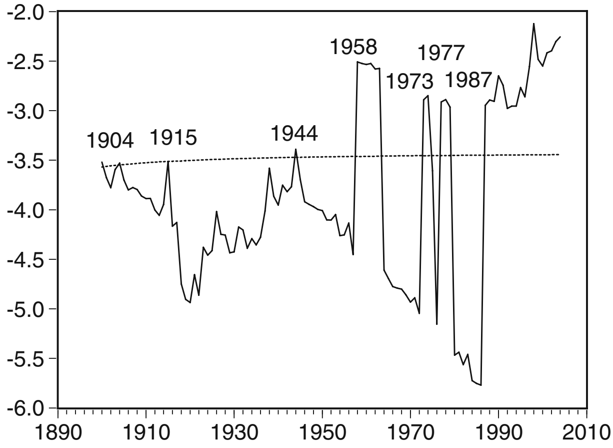
Five of the most commonly used unit root tests (Dickey and Fuller 1979; Said and Dickey 1984; Kwiatkowski et al. 1992; Elliott et al. 1996; Ng and Perron 2001) were applied to the global and hemispheric temperature series spanning over the 1870–2004 period (Jones and Moberg 2003; Jones et al. 2001; Folland et al. 2001a, b). Results show that all of them support the hypothesis of a unit root (see Standard unit root tests section in Appendix).

However, unit root tests are known to have little power to distinguish between trend-stationary and drifting unit root processes. Moreover, these tests are very sensitive to the specification of the deterministic component and to the number of lags added to account for serial correlation (Maddala and Kim 1998; Enders 2003; Hatanaka 1996).

Figure 1 shows the τ_τ statistic and the corresponding critical value (at 5% significance level) for a recursive Augmented Dickey Fuller (ADF) test (Dickey and Fuller 1979; Said and Dickey 1984) applied to the global temperature series starting with 30 observations and increasing sample size adding one observation at a time. τ_τ statistic values smaller than the critical values (MacKinnon 1996) imply the rejection of the null hypothesis of a unit root, while values greater than these values lead to the acceptance of the null. This is not intended to be a formal unit root testing procedure but just an illustrative and informative example. For most of the sample, the ADF test would have led to a rather strong rejection of the unit root hypothesis.[3] The dates corresponding to the sample size where the ADF test favors the unit

[3] Banerjee et al. (1992) proposed this type of recursive ADF tests for testing the null of a unit root against the alternative of a stationary process with a trend-break. The purpose of Fig. 1 is only to show that if we had estimated the ADF test each year since 1900, most of the time we would have rejected the null of a unit root. Nevertheless, if we use the critical values in Banerjee et al. ($t_{DF}^{\min} = -4.33$; $t_{DF}^{\max} = -1.99$, 5%), the null of a unit root is strongly rejected in favor of trend stationarity with a break.

Fig. 1 Recursive ADF test for an increasing sample size of the observed global temperature series



root hypothesis are: 1904, 1915, 1944, 1958, 1973, 1977 and 1987. All these years correspond to El Niño episodes, which are normally associated with large positive variations in temperature (note that the largest value of the τ_τ statistic corresponds to 1998, one of the strongest El Niños). On the other hand, this figure also shows that after 1987 the ADF permanently indicates the presence of a unit root.[4] This behavior suggests that this test might not be able to reject the hypothesis of a unit root in the global temperature series because of the presence of possible structural changes and large realizations of the process.

Perron (1989) showed that in the presence of a shift in the trend function, the first order autorregresive coefficient is highly biased towards unity. In this case, the unit root null is hardly rejected even if the series is composed of a trend (with a break) and *i.i.d.* disturbances. Furthermore, when the break occurs in the slope of the trend function, unit root tests are not consistent and the null hypothesis cannot be rejected even asymptotically.

Consequently, several tests have been developed for testing the null of difference-stationary against the alternative of trend-stationary with permanent shocks in the trend function (Perron 1989, 1997; Zivot and Andrews 1992; Lumsdaine and Papell 1997; Harvey and Mills 2002; Banerjee et al. 1992; Kim and Perron 2007). The tests of Perron (1997), Zivot and Andrews (1992) and Kim and Perron (2007) were applied to the global and hemispheric temperature series (see Methods section in Appendix). The Zivot and Andrews relies on the assumption of endogeneity of the shock which consists in that the break in the trend function is interpreted as a realization from the tail of the distribution of the underlying data-generating process. Recent literature has shown that, as a consequence of this assumption, a rejection of the null of a unit root could be due to the presence of a break even if the unit root is present. Moreover, if the noise component has a unit root and a break in the trend function

[4]Note that the recursive ADF test in Fig. 1 does not (and is not meant to) provide an estimate of the break date in the trend function.

occurs, then the Zivot-Andrews tests often diverge or are not invariant to break parameters (Vogelsang and Perron 1998; Kim and Perron 2007). For these reasons, the results of this test are omitted but are available upon request.

These problems are not present in the Perron (1997) test when the break date is selected by minimizing the sum of the squared residuals. Furthermore, Kim and Perron (2007) provide a testing procedure that allows a break in the trend function under both the null and the alternative hypotheses and that has the same limit distribution as if the break date was known. These tests have greater power, maintain the correct size and simulation experiments have shown that they offer an improvement over other commonly used methods in small samples.

Perron (1989, 1997) and Kim and Perron (2007) assume an exogenous break in both the null and the alternative hypotheses. In this case the null hypothesis consists of a random walk in temperature series with a structural change produced by an external forcing factor against the alternative of a trend-stationary process with a structural change also produced by an external forcing factor which has changed the long-run path of temperature.

Table 1 presents the results of applying the Perron and the Kim and Perron tests. As can be seen in this table, once a break in the trend function is allowed, the hypothesis of a unit root in global and hemispheric temperatures is strongly rejected by both tests. The estimations of the first order autoregressive coefficient are quite far from the unity: 0.50 for the global temperature series, 0.48 for the Northern Hemisphere and 0.42 for the Southern Hemisphere. Assuming a unit root would have attributed to much persistence to temperature variability (as a matter of fact, an infinite persistence). The response to a one-time shock of one standard deviation for global and hemispheric temperatures was estimated using an impulse–response function. A shock of such magnitude in average will produce an accumulated long-run response of 0.19°C in the global temperature (G), 0.23°C in the Northern Hemisphere temperature (NH) and 0.16°C in the Southern Hemisphere (SH). As opposed to the unit root hypothesis, most of these shocks (92% for SH, 89% for NH and 88% for G) would dissipate in an average of 3 years.

Figure 2 shows global, Northern and Southern hemisphere temperature series, and their corresponding linear (with break) and nonlinear trends (Hodrick and Prescott 1997). The linear trend specification allows large natural variations to be modeled as part of the noise function and therefore not assuming a deterministic function for large natural variations as opposed to Harvey and Mills (2000, 2002).

Perron (1997) test shows that the estimated break dates are: 1977 for G, 1985 for NH and 1911 for SH. Given that this test adopts a two step procedure[5] that provides a consistent estimate of the break date (or more precisely of the break fraction), it is possible to construct a confidence interval for it using the methodology in Perron and Zhu (2005). The 95% confidence intervals for the break date are: 1972–1982 for G, 1981–1989 for NH and 1904–1918.

The estimated trend coefficients before the structural change for G and NH are positive (0.0035°C and 0.0036°C per year respectively) and highly significant. The post-break slopes of G and NH are notably higher: 0.018°C/year and 0.030°C/year respectively. Note that the magnitude of the post-break slope in G is 5 times the

[5] See the Methods section in Appendix.

Table 1 Tests for a unit root with a one-time break in the trend function

Series	T_b	k	$\hat{\mu}$	$t_{\hat{\mu}}$	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\gamma}$	$t_{\hat{\gamma}}$	$\hat{\alpha}$	$t_{\hat{\alpha}}$	$S(\hat{e})$	$t_\alpha(\hat{\lambda}_{tr}^{AO})$
Globe	1977	0	−0.4134	**−19.766**	0.0035	**10.844**	0.0142	**7.8539**	0.4987	−6.6978a	0.0960	−5.7307A
Nhem	1985	0	−0.3813	**−15.079**	0.0036	**9.7751**	0.0260	**7.8327**	0.4775	−6.9027a	0.1207	−5.8863A
Shem	1911	0	−0.3402	**−11.650**	−0.0018	**−1.9200**	0.0088	**7.3759**	0.4233	−7.3492a	0.0936	−6.9026A

The regression model for the unit root tests is defined in Eqs. 15 and 16 of the "Methods" section. The symbols used above have meanings as follows: T_b is the estimated time of the break; k is the number of lagged differences added to correct for serial autocorrelation; $S(\hat{e})$ is the standard error of 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. Bold numbers denote statistical significance at 5% levels. $\hat{\alpha}$ is the first order autoregressive coefficient and $t_{\hat{\alpha}}$ is the unit root test statistic. a, b, c, d denote statistical significance at the 1%, 2.5%, 5% and 10% respectively (Perron (1997) critical values). $t_\alpha(\hat{\lambda}_{tr}^{AO})$ is the unit root test statistic corresponding to the Kim and Perron (2007) test. A, B, C, D denotes statistical significance at the 1%, 2.5%, 5% and 10% respectively (Perron and Vogelsang (1993) critical values Table 1).

The existence of a break in the slope of the trend function was tested using the Exp–Wald statistic of the Andrews–Ploberger test correcting for first order autocorrelation and the critical values shown in Perron and Yabu (2007). The Exp–Wald statistic values are: 3.59 (globe), 4.24 (nhem) and 3.53 (shem). All are significant at 5% levels with a 5% trimming

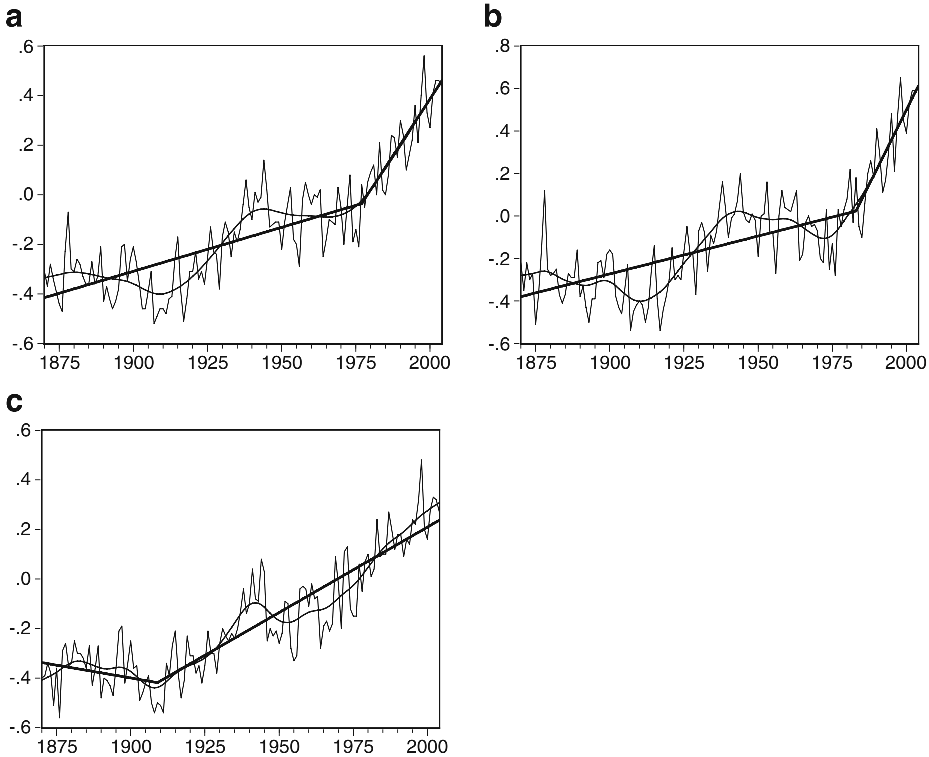


Fig. 2 Panels **a**, **b** and **c** show the observed global, Northern and Southern Hemisphere temperature series and their linear and non-linear trends. The estimated dates of the one-time break in the linear trends are 1977, 1985 and 1911 respectively

value of its pre-break slope, and that in the case of NH, the magnitude of the post-break slope is more than 8 times the value of its pre-break slope. These observed temperature series provide evidence of a "two stages" climate change: a moderate warming for most of the twentieth century, followed by a rapid and of a much higher magnitude in the last two to three decades.

On the other hand, the SH shows a slight cooling trend until 1910 and a moderate warming after this date. Since a large part of the Southern Hemisphere corresponds to ocean, the delay in the beginning of the warming trend in comparison to the G and NH temperature series could be interpreted as the slow reaction of the oceans to changes in forcing factors given their known high heat capacity. In this case, the moderate warming trend might correspond to the first warming stage shown by the other two series and therefore a future change in the slope of a possible higher magnitude could be expected in the future.

The tests applied in this paper provide strong evidence against the hypothesis of the twentieth century warming being a realization of an externally unforced global and hemispheric temperature data-generating process, in favor of a trend-stationary process with a structural break in its rate of growth caused by an external forcing factor. The rejection of the null on these tests gives strong evidence against the unit root being an adequate representation of the data-generating process of global and hemispheric temperature series and in favor of a trend-stationary process with an exogenous shock with permanent effects. Perron and Kim and Perron[6] tests indicate that the only "permanent shock" in global and hemispheric temperatures corresponds to a change in the rate of growth caused by an external forcing.

Misspecification tests[7] provide evidence of normal-linear-homoskedastic stationary errors (Andreou and Spanos 2003; Spanos and Mcguirk 2002). In particular it is important to notice that there is no evidence of changes in the variance of the process (the errors are homoskedastic). Observed climate change in global and hemispheric annual temperatures seems to have manifested itself as a change-in-mean, constant variance, phenomenon. This has obvious and important implications for risk assessment and temperature forecasting.

Interpreted narrowly, the specification of a deterministic trend function with one-time break is certainly restrictive and unattractive (Banerjee et al. 1992), particularly for forecast purposes, but if interpreted broadly, it shows that there are few large variations in the key forcing factors that can change the long run path of global and hemispheric temperature series at least in the time lapse of the sample.

Moreover, as explained below, our interpretation of the trend function is based on Perron (1989) where the trend function and its changes are not assumed to be deterministic. Consider the following parameterization:

$$y_t = \eta_t + Z_t \tag{6}$$

$$\eta_t = \mu_t + \beta_t t \tag{7}$$

where $A(L)Z_t = B(L)e_t$; $e_t \sim i.i.d.(0, \sigma^2)$; $\mu_t = \mu_{t-1} + V(L)\, v_t$; $\beta_t = \beta_{t-1} + W(L)\, w_t$

The process is composed as a trend function η_t and a noise function Z_t. The intercept and slope of the trend function are considered integrated stochastic processes with $W(L)$ and $V(L)$ satisfying the stationarity and invertibility conditions. However, the occurrence of the shocks w_t and v_t are rare in comparison with the sequence of innovations $\{e_t\}$. That is, the coefficients of the trend function (i.e. the long-run path of global and hemispheric temperatures or climate) are determined by realizations of w_t and v_t that are (large) changes in key external forcing factors such as Earth orbit changes, solar irradiance, and greenhouse gases concentrations that hardly occur in decadal and century time scales. Non-zero realizations of the process w_t occurred in 1977, 1985 and 1911 in G, NH and SH respectively. The sequence of innovations $\{e_t\}$ represents the short-term variations.

As shown above, the trend function and its changes are not considered to be deterministic. As stated in Perron (1989) "the exogeneity assumption about the changes in the trend function is a device that allows taking these shocks out of the noise function into the trend function without specific modeling of the stochastic behavior of μ_t and β_t". The purpose of unit root with a structural change tests is to "test whether Z_t is an integrated process or not, i.e. to test whether the shocks $\{e_t\}$ have persistent effects that do not vanish over a long horizon" (Perron 1989)."

[6]The Zivot-Andrews test also lead to the rejection of the unit root null.

[7]Test results are available upon request.

A review of recent literature on climate change provides arguments for entertaining the hypothesis of a possible anthropogenic origin of the structural breaks found in the trend functions of global and hemispheric temperatures. For example, the Summary for Policy Makers of the Working Group I of the IPCC's Forth Assessment Report (IPCC 2007) states that "most of the increase in global average temperatures since the mid-twentieth century is very likely (90% confidence) due to the observed increase in anthropogenic greenhouse gas concentrations". Furthermore, recent paleoclimate studies (Mann et al. 1998, 1999; Jones and Mann 2004; Mann and Jones 2003; Crowley 2000) conclude the most plausible source of the late twentieth century warming is anthropogenic forcing. For example, Jones and Mann (2004) analyzed climate changes over the past 2 millennia concluding that natural factors appear to explain most of the changes in surface temperature up to the nineteenth century, but that only anthropogenic forcing can explain the unprecedented warming of the late twentieth century. Their analysis of forcing factors shows that:

- Solar irradiance variability was relatively small compared with volcanic and greenhouse forcing
- The dearth of volcanic eruptions from 1915 to 1965 has been pointed as a potential explanation of the unusual warming of the twentieth century. Nevertheless, in the context of longer-term climate changes, the period from 1000 to 1500 experienced relatively fewer eruptions in comparison to the 1500–2000 period.
- There is a fairly robust positive trend in greenhouse gas concentrations. For example, $CO2$ concentrations have increased from 275 ppm in the pre-industrialization period to 320 ppm in the mid twentieth century.
- Land use change forcing is estimated to have produced a 0.4°C net cooling effect during the past few centuries.

In the same paper, Jones and Mann also reported that the empirical evidence provided by climate change proxies over the past two millennia in combination with climate models simulations indicates that, while solar and volcanic forcing played dominant roles among the potential causes of climate variability, these forcing factors are unable to explain the dramatic warming of the late twentieth century. Moreover, according to these authors, only anthropogenic influences are able to explain such warming given that natural forcing factors would favor a slight cooling over this period.

4 Conclusions

Results show strong statistical evidence against a stochastic trend in global and hemispheric temperatures. The data-generating process of the temperature series analyzed can be assumed as trend stationary with only one permanent shock occurring at different dates and of different magnitudes. Cointegration, statistical tests and inferences that are constructed assuming that temperatures are unit root processes are not reliable.

For the global and Northern Hemisphere temperatures results show a "two stages" warming trend, the first a moderate one and the other a rapid and of much higher magnitude. The warming in the Southern Hemisphere began later than in the Northern Hemisphere and in global temperatures, possibly because the oceans' storing heat capacity. The behavior shown by G and NH series could indicate that the SH has not yet fully responded to the changes in external forcing factors and therefore a "second stage" warming could be expected.

The time series properties of global and hemispheric temperatures reveal that climate change has manifested itself as a "change-in-the-mean phenomenon", variability has not increased (or decreased) with climate change. This is of outmost importance for modeling and forecasting future temperature and for risk assessment.

Although the objective of this study is to analyze the time series properties of global and hemispheric temperatures and not to establish direct climate change attribution, important inferences can be obtained regarding this issue. The statistical tests performed on the series provide evidence about the exogenous nature of the shock and reinforce the findings of recent paleoclimate studies (Mann et al. 1998, 1999; Jones and Mann 2004; Mann and Jones 2003; Crowley 2000) in that the twentieth century warming cannot be interpreted as a realization of natural temperature variability. According to those studies, the most plausible source of the twentieth century warming is anthropogenic forcing.

In terms of the Article 2 of the Framework Convention on Climate Change it can be argued that significant (dangerous) anthropogenic interference with the climate system has already occurred. Moreover, the post-break global temperature trend is very close to the "the tolerable window" (Bruckner and Schellnhuber 1999) of 0.2°C/decade originally proposed by the German Advisory Council on Global Change (WBGU) and the late trend shown by the Northern Hemisphere has already surpassed it. Quoting Tony Blair (Joachim et al. 2006): "what level of greenhouse gases in the atmosphere is self-evidently too much?", it looks like we might have already exceeded it.

Acknowledgements The authors would like to thank Victor Guerrero and Manuel Mendoza from the Statistics Department at ITAM, and an anonymous reviewer for their valuable comments. Gay and Estrada acknowledge financial support for this work from the Consejo Nacional de Ciencia y Tecnología under grant CONACYT-SEMARNAT-23836.

Appendix

Standard unit root tests

Table 2 Augmented Dickey–Fuller test

Series	Model A	Model B	Model C
Globe	−2.257 (3)	0.414 (5)	−0.781 (3)
ΔGlobe	**−8.071 (4)**	**−11.024 (2)**	**−10.900 (2)**
Nhem	−2.026 (3)	−0.020 (3)	−0.577 (3)
ΔNhem	**−11.775 (2)**	**−11.666 (2)**	**−11.572 (2)**
Shem	−3.402 (2)	−0.502 (3)	−1.167 (3)
ΔShem	**−10.456 (2)**	**−10.435 (2)**	**−10.347 (2)**

Model A includes a constant and a linear trend; Model B includes a constant; Model C includes no deterministic variables. Figures in bold indicate that the statistic is significant at the 5% level. Lag length is given in parentheses. The critical values are: Model A −4.029 (1%), −3.444 (5%), −3.147 (10%); Model B −3.481 (1%), −2.883 (5%), −2.578 (10%); Model C −2.582 (1%), −1.943 (5%), −1.615 (10%). Lag length selection: Akaike Information Criterion

Table 3 Elliot, Rothenberg, and Stock Point Optimal (ERS) test

Series	Model A	Model B
Globe	43.112 (9)	22.454 (3)
ΔGlobe	**1.501 (0)**	**0.478 (0)**
Nhem	59.895 (9)	16.945 (3)
ΔNhem	**1.688 (0)**	**0.689 (0)**
Shem	16.855 (9)	21.496 (3)
ΔShem	**1.462 (0)**	**0.401 (2)**

Model A includes a constant and a linear trend; Model B includes a constant. Lag length is given in parentheses. Figures in bold indicate that the statistic is significant at the 5% level. The critical values are: Model A 4.186 (1%), 5.647 (5%), 6.814 (10%); Model B 1.936 (1%), 3.131 (5%), 4.226 (10%). Lag length selection: Modified Akaike Information Criterion (Ng and Perron 2001)

Table 4 Elliott–Rothenberg–Stock DF-GLS test

Series	Model A	Model B
Globe	−1.297 (5)	0.507 (3)
ΔGlobe	−2.333 (11)	−0.561 (11)
Nhem	−1.603 (3)	0.224 (3)
ΔNhem	−2.109 (10)	−0.041 (11)
Shem	−1.455 (9)	0.215 (3)
ΔShem	**−14.693 (0)**	**−14.764 (0)**

Model A includes a constant and a linear trend; Model B includes a constant. Lag length is given in parentheses. Figures in bold indicate that the statistic is significant at the 5% level. The critical values are: Model A −3.542 (1%), −2.999 (5%), −2.709 (10%); Model B −2.583 (1%), −1.943 (5%), −1.615 (10%). Lag length selection: Modified Akaike Information Criterion (Ng and Perron 2001)

Table 5 KPSS test

Series	Model A	Model B
Globe	**0.194**	**1.224**
ΔGlobe	0.063	0.186
Nhem	**0.154**	**1.133**
ΔNhem	0.084	0.261
Shem	**0.233**	**1.270**
ΔShem	0.032	0.076

Model A includes a constant and a linear trend; Model B includes a constant. Lag length is given in parentheses. Figures in bold indicate that the statistic is significant at the 5% level. The critical values are: Model A 0.216 (1%), 0.146 (5%), 0.119 (10%); Model B 0.739 (1%), 0.463 (5%), 0.347 (10%). Bandwidth selection: Newey–West with Bartlett kernel

Table 6 Ng–Perron Modified unit root tests

Series	MZa	MZt	MSB	MPT
Globe				
Model A	−4.208 (5)	−1.267 (5)	0.301 (5)	19.940 (5)
Model B	1.117 (3)	0.510 (3)	0.457 (3)	20.415 (3)
ΔGlobe				
Model A	**−61.405 (0)**	**−5.538 (0)**	**0.090 (0)**	**1.500 (0)**
Model B	**−54.002 (0)**	**−5.196 (0)**	**0.096 (0)**	**0.456 (0)**
Nhem				
Model A	−6.061 (3)	−1.531 (3)	0.253 (3)	14.889 (3)
Model B	0.788 (3)	0.303 (3)	0.384 (3)	15.799 (3)
ΔNhem				
Model A	**−55.581 (0)**	**−5.270 (0)**	**0.095 (0)**	**1.649 (0)**
Model B	**−42.046 (0)**	**−4.576 (0)**	**0.109 (0)**	**0.608 (0)**
Shem				
Model A	−5.246 (9)	−1.558 (9)	0.297 (9)	17.149 (9)
Model B	0.232 (3)	0.112 (3)	0.481 (3)	19.040 (3)
ΔShem				
Model A	**−65.749 (0)**	**−5.730 (0)**	**0.087 (0)**	**1.403 (0)**
Model B	**−69.643 (0)**	**−5.896 (0)**	**0.085 (0)**	**0.364 (0)**

Model A includes a constant and a linear trend; Model B includes a constant. Lag length is given in parentheses. Figures in bold indicate that the statistic is significant at the 5% level. The critical values are: Model A: MZa −23.800 (1%), −17.300 (5%), −14.200 (10%); MZt −3.420 (1%), −2.910 (5%), −2.620 (10%); MSB 0.143 (1%), 0.168 (5%), 0.185 (10%); MPT 4.030 (1%), 5.480 (5%), −6.670 (10%); Model B: MZa −13.800 (1%), −8.100 (5%), −5.700 (10%); MZt −2.580 (1%), −1.980 (5%), −1.620 (10%); MSB 0.174 (1%), 0.233 (5%), 0.275 (10%); MPT 1.780 (1%), 3.170 (5%), −4.450 (10%). Lag length selection: Modified Akaike Information Criterion (Ng and Perron 2001). The modification shown in Perron and Qu (2007) was applied to ΔGlobe and ΔNhem.

Methods

Perron (1989, 1997) considers three different models under the null hypothesis: a "crash" model that allows for an exogenous change in the level of the series (Model (A)); a "changing growth" model that permits an exogenous change in the rate of growth (Model (B)); and one model that allows for both of these changes to occur at the same time (Model (C)). The null hypotheses are parameterized as follows:

$$\text{Model }(A)\; y_t = \mu + dD\,(TB)_t + y_{t-1} + e_t, \tag{8}$$

$$\text{Model }(\mathrm{B})\; y_t = \mu_1 + y_{t-1} + (\mu_2 - \mu_1)\, DU_t + e_t, \tag{9}$$

$$\text{Model }(\mathrm{C})\; y_t = \mu_1 + y_{t-1} + dD\,(TB)_t + (\mu_2 - \mu_1)\, DU_t + e_t, \tag{10}$$

where

$D(TB)_t = 1$ if $t = T_b + 1, 0$ otherwise;
$DU_t = 1$ if $t > T_b, 0$ otherwise;
and $A(L)e_t = B(L)v_t$,
$v_t \sim$ i.i.d. $(0, \sigma^2)$, with A(L) and B(L) being the pth and qth order polynomials in the lag operator.

The models considered under the alternative hypothesis are:

$$\text{Model } (A)\; y_t = \mu_1 + \beta t + (\mu_2 - \mu_1)\, DU_t + e_t, \tag{11}$$

$$\text{Model (B)}\,,\; y_t = \mu + \beta_1 t + (\beta_2 - \beta_1)\, DT_t^* + e_t \tag{12}$$

$$\text{Model (C)}\; y_t = \mu_1 +_t \beta_1 t + (\mu_2 - \mu_1)\, DU_t + (\beta_2 - \beta_1)\, DT + e_t, \tag{13}$$

where

$DT_t^* = t - T_b\,,\; DT_t = t$ if $t > T_b, 0$ otherwise. T_b refers to the time of the break.

For our study we proposed Model (B) for global and Northern and Southern hemispheric temperature series. To carry out the testing procedure, the following regressions were estimated using ordinary least squares (OLS):

$$y_t = \tilde{\mu} + \tilde{\beta} t + \tilde{\gamma}\, DT_t^* + \tilde{y}_t \tag{14}$$

$$\tilde{y}_t = \tilde{\alpha}\, \tilde{y}_{t-1} + \sum_{i=1}^{k} \tilde{c}_i \Delta \tilde{y}_{t-i} + \tilde{e}_t \tag{15}$$

That is, the raw data is first detrended and then an ADF (Dickey and Fuller 1979; Said and Dickey 1984) type regression is estimated. This is an additive outlier approach and implies that the change in the slope of the trend function occurs rapidly (abrupt change). The break date was selected by minimizing the sum of squared residuals of regression (14). The lagged differences of y_t are included to correct for autocorrelation (parametrically). Determining the number of extra regressors of $\Delta \tilde{y}_{t-i}$ (that is, k) is an important issue because if too many extra regressors will decrease the power of the test, while including too few may have important size effects. Several data dependent methods are discussed in econometric literature (Perron 1997; Ng and Perron 2001, 2005) for selecting the value of k. For our study we followed the approach of Spanos and Mcguirk (2002) and Andreou and Spanos (2003) and construct statistically adequate models that ensure "the empirical validity of the probabilistic assumptions underlying a statistical model". According to Andreou and Spanos (2003), statistical adequacy provides sufficient condition for statistical inference because it presupposes the validity of the model specification assumptions. Our results are quite robust to alternative methods for choosing the lag length such as the Akaike Information Criterion (AIC), Schwarz Information Criterion (BIC) and the Hannan-Quinn criterion (HQ). A *k-max* of 3 lags was chosen; results are available upon request.

The Kim and Perron (2007) test procedure consists in the following steps:

a) Given an estimate of the break date $\hat{T} = \hat{\lambda} T$, set a window around it and define T_l and T_h;
b) Create a new data set $\{y^n\}$ by removing the data from T_l +1 to T_h, and shifting down the data after the window by $S(T) = y_{T_h} - y_{T_l}$;

c) Perform the unit root test corresponding to Model (B) using the break date T_l and compute the $t_\alpha\left(\hat{\lambda}_{tr}^{AO}\right)$ statistic from the following regression (16):

$$\tilde{y}_t^n = \tilde{\alpha}\tilde{y}_{t-1}^n + \sum_{i=1}^{k} \tilde{c}_i \Delta \tilde{y}_{t-i}^n + \tilde{e}_t \tag{16}$$

where $\hat{\lambda}_{tr} = T_l \big/ T_r$ and $\hat{y}_t^n$ is the detrended value of y^n.

Estimations of the break date and unit root tests were obtained using the Pierre Perron and Serena Ng RATS procedure, and an Eviews program developed by the authors.

References

Andreou A, Spanos A (2003) Statistical adequacy and the testing of trend versus difference stationary. Econom Rev 22:217–237

Banerjee A, Lumsdaine RL, Stock JH (1992) Recursive and sequential tests of the unit root and trend break hypotheses: theory and international evidence. JBES 10(3):271–287

Bloomfield P (1992) Trend in global temperature. Clim Change 21:1–16

Bruckner T, Schellnhuber HJ (1999) Climate change protection: the tolerable windows approach. (IPTS Report 34)

Crowley TJ (2000) Causes of climate change over the past 1000 years. Science 289:270–277

Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc 74:427–431

Elliott G, Rothenberg TJ, Stock JH (1996) Efficient tests for an autoregressive unit root. Econometrica 64:813–836

Enders W (2003) Applied econometric time series. Wiley, New York

Folland CK et al (1992) Observed climate variability and change. In: Houghton JT, Callander BA, Varney SK (eds) Climate change 1992: the supplementary report to the IPCC scientific assessment. Cambridge University Press, Cambridge

Folland CK et al (2001a) Global temperature change and its uncertainties since 1861. Geophys Res Lett 28:2621–2624

Folland CK et al (2001b) Observed climate variability and change. In: Houghton JT et al (eds) Climate change 2001: the scientific basis. Cambridge Univ Press, Cambridge, pp 99–181

Galbraith J, Green C (1992) Inference about trends in global temperature data. Clim Change 22: 209–221

Gay C, Estrada F, Conde C (2007) Some implications of time series analysis for describing climatologic conditions and for forecasting. An illustrative case: Veracruz, Mexico. Atmosfera 20(2):147–170

Hansen J, Lebedeff S (1987) Global trends of measured surface air temperature. J Geophys Res 92:13345–13372

Harvey DI, Mills TC (2000) Detecting climate change using smooth transitions. (Economic Research Paper 00/21, Department of Economics, Loughborough University); (available at http://www.lboro.ac.uk/departments/ec/Research/Discussion%20Papers%202000/Research%20Papers%202000/erp00–21.pdf)

Harvey DI, Mills TC (2002) Unit roots and double smooth transitions. J Appl Stat 29(5):675–683

Hatanaka M (1996) Time-series-based econometrics. Unit roots and cointegration. Oxford University Press, Oxford

Hodrick RJ, Prescott EC (1997) Postwar U.S. business cycles: an empirical investigation. J Money Credit Bank 29:1–16

IPCC (2001) Climate change 2001: synthesis report 2001. In: Watson RT, the Core Writing Team (eds) Contribution of working group I, II, and III to the third assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, 397 pp

IPCC (2007) Climate change 2007: the physical science basis. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, 996 pp

Joachim H, Schnellnhuber HJ et al (ed) (2006) Avoiding dangerous climate change. Cambridge University Press, Cambridge

Jones PD, Mann ME (2004) Climate over past millennia. Rev Geophys 42:RG2002. doi:10.1029/2003RG000143

Jones PD, Moberg A (2003) Hemispheric and large-scale surface air temperature variations: an extensive revision and an update to 2001. J Clim 16:206–223

Jones PD et al (2001) Adjusting for sampling density in grid-box land and ocean surface temperature time series. J Geophys Res 106:3371–3380

Kärner O (1996) Global temperature deviations as a random walk. J Climate 9:656–658

Kaufmann RK, Stern DI (1997) Evidence for human influence on climate from hemispheric temperature relations. Nature 388:39–44

Kaufmann RK, Kauppi H, Stock JH (2006a) Emissions, concentrations, & temperature: a time series analysis. Clim Change 77:249–278. doi:10.1007/s10584-006-9062-1

Kaufmann RK, Kauppi H, Stock JH (2006b) The relationship between radiative forcing and temperature: what do statistical analyses of the instrumental temperature record measure? Clim Change 77:279–289. doi:10.1007/s10584-006-9063-0

Kim D, Perron P (2007) Unit root tests allowing for a break in the trend function under both the null and the alternative hypotheses. Unpublished manuscript. Department of Economics, Boston University

Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. J Econom 54:159–178

Lumsdaine R, Papell DH (1997) Multiple trend breaks and the unit-root hypothesis. REStat 79(2):212–218

MacKinnon JG (1996) Numerical distribution functions for unit root and cointegration tests. J Appl Econ 11:601–618

Maddala GS, Kim IM (1998) Unit roots, cointegration and structural change. Themes in Modern Econometrics, Cambridge Univ Press

Mann ME, Jones PD (2003) Global surface temperature over the past two millennia. Geophys Res Lett 30(15):1820. doi:10.1029/2003GL017814

Mann ME, Bradley RS, Hughes MK (1998) Global scale temperature patterns and climate forcing over the past six centuries. Nature 392:779–787

Mann ME, Bradley RS, Hughes MK (1999) Northern Hemisphere temperatures during the past millennium: inferences, uncertainties and limitations. Geophys Res Lett 26:759–762

Nelson CR, Plosser CI (1982) Trends and random walks in macroeconomic time series: some evidence and implications. J Monet Econ 10:139–162

Ng S, Perron P (2001) Lag length selection and the construction of unit root tests with good size and power. Econometrica 69(6):1519–1554

Ng S, Perron P (2005) A note on the selection of time series models. OBES, Department of Economics, University of Oxford, vol 67, no 1, pp 115–134, 02

Perron P (1989) The great crash, the oil price shock, and the unit root hypothesis. Econometrica 99:1361–1401

Perron P, Vogelsang T (1993) Erratum: the great cash, the oil price shock and the unit root hypothesis. Econometrica 61:248–249

Perron P (1997) Further evidence on breaking trend functions in macroeconomic variables. J Econom 80(2):355–385

Perron P, Qu Z (2007) A simple modification to improve the finite sample properties of Ng and Perron's unit root tests. Econ Lett 94:12–19

Perron P, Yabu T (2007) Testing for shifts in trend with an integrated or stationary noise component. Forthcoming in JBES. Available at http://people.bu.edu/perron/papers/breakpaper.pdf

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

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

Schneider SH (2002) Can we estimate the likelihood of climatic changes at 2100? Clim Change 52(4):441–451

Spanos A, Mcguirk A (2002) Where do statistical models come from? The problem of specification uncertainty in empirical modeling. (Virginia Tech working paper) (available at http://www.econ.ucy.ac.cy/seminars/Spanos.pdf)

Stern DI, Kaufmann RK (1997) Time series properties of global climate variables: detection and attribution of climate change. (Working Papers in Ecological Economics. The Australian National University, Center for Resource and Environmental Studies Ecological Economics Programme) (available at http://eprints.anu.edu.au/archive/00000665/00/eep9702.pdf)

Stern DI, Kaufmann RK (1999) 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)

Stock JH (1994) Unit roots, structural breaks, and trends. In: Engle R, McFadden D (eds) Handbook of econometrics, vol IV, pp 2739–2841. North-Holland, New York

Vogelsang TJ, Perron P (1998) Additional tests for a unit root allowing the possibility of breaks in the trend function. Int Econ Rev 39:1073–1100

Woodward WA, Gray HL (1993) Global warming and the problem of testing for trend in time series data. J Clim 6:953–962

Woodward WA, Gray HL (1995) Selecting a model for detecting the presence of a trend. J Clim 8:1929–1937

Zheng X, Basher RE (1999) Structural time series models and trend detection in global and regional temperature series. J Clim 12:2347–2358

Zivot E, Andrews D (1992) Further evidence on the Great Crash, the oil price shock, and the unit root hypothesis. JBES 10(3):251–270