Universidad de Navarra

Facultad de Ciencias Económicas y Empresariales

Working Paper nº 09/08

Time trend estimation with breaks in temperature time series

Luis A. Gil-Alana
Facultad de Ciencias Económicas y Empresariales
Universidad de Navarra

Time trend estimation with breaks in temperature time series
Luis A. Gil-Alana
Working Paper No.09/08
November 2008

ABSTRACT

This paper deals with the modelling of the global and northern and southern hemispheric anomaly temperature time series using a novel technique based on segmented trends and fractional integration. We use a procedure that permits us to estimate linear time trends and orders of integration at various subsamples, where the periods for the changing trends are endogenously determined by the model. Moreover, we use a non-parametric approach (Bloomfield, 1973) for modelling the I(0) deviation term. The results show that the three series (global, northern and southern temperatures) can be well described in terms of fractional integration with the orders of integration around 0.5 in the three cases. The coefficients associated to the time trends are statistically significant in all subsamples for the three series, especially during the final part of the sample, giving then some support to the global warming theories.

Luis A. Gil-Alana
Universidad de Navarra
Dpto. Métodos Cuantitativos
Campus Universitario
31080 Pamplona
alana@unav.es

1. Introduction

Global and hemispheric temperatures have increased during the last one hundred and fifty years. Due to the social interest of the climate change issue, it is not uncommon to find papers analysing a single time series of a scalar climatic variable to check if it contains a trend that might indicate how much and in what direction temperatures are changing. On the other hand, it is a well known stylized fact that temperatures are time-dependent. Mathematically, there exist different ways of modelling that dependence and a key issue here is to determine if the series is stationary (around a linear time trend) or not.[1] To motivate our discussion, let us suppose that y_t is the temperature series observed at a given location at time t. We can consider a linear time trend model (of the form as in equation (6) described below) and, if we suppose that the detrended series (x_t) is stationary, the AutoRegressive of order 1 (AR(1)) process is one of the simplest models to describe such dependence. The AR(1) model is described as:

$$x_t = \rho x_{t-1} + \varepsilon_t, \quad t = 1, 2, \ldots, \tag{1}$$

where $|\rho| < 1$ and ε_t is white noise. This model has been widely employed in the climatological community because of its relation with the stochastic first order differential equation. On the other hand, if we believe the series is nonstationary (with or without a linear trend), we can take ρ in equation (1) equal to 1, and the process is then said to be integrated of order 1 (and denoted by $x_t \sim I(1)$). In other words, x_t is then nonstationary but its first differences $(1 - L)x_t = x_t - x_{t-1}$ are stationary, and statistical inference must be relied on the differenced process. In fact, the differenced process can be more general than white noise and it may contain some other type of dependence that may be described through an AR process. Rigorously speaking, x_t is said to be I(1) if:

$$(1 - L) x_t = u_t, \quad t = 1, 2, ..., \tag{2}$$

where u_t is I(0) defined as a covariance stationary process where the infinite sum of the autocovariances is finite.[2] In such a case, the series is said to contain a stochastic trend or a unit root. Evidence of this type of model in temperature time series is found in Woodward and Gray (1995), Stern and Kaufmann (2000), Kaufmann and Stern (2002), Kaufmann et al. (2006), etc.[3]

During the last twenty years other approaches have become very popular when modelling the nonstationarity (or even the stationary part) in time series. In particular, the number of differences required to achieve an I(0) process may not necessarily be an integer value (usually 1) but may also be a real value between 0 and 1. In such a case, the process is said to be fractionally integrated or integrated of order d (I(d)) where d may be a real value. That is, x_t is said to be I(d) if

$$(1 - L)^d x_t = u_t, \quad t = 1, 2, ..., \tag{3}$$

with I(0) u_t. Note that the polynomial in the left-hand-side in (3) can be expressed in terms of its Binomial expansion, such that, for all real d,

$$(1 - L)^d = \sum_{j=0}^{\infty} \psi_j L^j = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j L^j = 1 - d\, L + \frac{d(d-1)}{2} L^2 - \dots . \tag{4}$$

and then equation (3) can be written as:

$$x_t = d\, x_{t-1} - \frac{d(d-1)}{2} x_{t-2} + \dots + u_t. \tag{5}$$

Thus, if d is an integer value, x_t will be a function of a finite number of past observations, while if d is real, x_t depends strongly upon values of the time series far away in the past.

(See, e.g., Granger and Ding, 1996; Dueker and Asea, 1998). Moreover, higher the d is, the higher will be the level of association between the observations. The differencing parameter d plays a crucial role from a statistical viewpoint. Thus, if $d \in (0, 0.5)$, the series is covariance stationary and mean-reverting, with shocks disappearing in the long run; if $d \in [0.5, 1)$, the series is no longer stationary but still mean-reverting, while $d \geq 1$ means nonstationarity and non-mean-reversion. It is therefore crucial to examine if d is smaller than or equal to or higher than 1. Comprehensive surveys of fractionally integrated models can be found in Robinson (1994, 2003), Beran (1994), Baillie (1996) and Doukhan, Oppenheim and Taqqu (2003) among others, and examples of this type of model in meteorological time series data are the papers of Bloomfield (1992); Smith (1993); Lewis and Ray (1997); Pethkar and Selvam (1997), Koscielny-Bunde et al. (1998); Pelletier and Turcotte (1999); Percival et al. (2004); Maraun et al. (2004) and Gil-Alana (2003, 2005).

As mentioned above, it is also a well-known stylized fact that temperatures have increased across time, especially during the XXth century. The standard way of modelling that behaviour is to assume a linear function of time of form:

$$y_t = \beta_0 + \beta_1 t + x_t, \qquad t = 1, 2, \ldots, \tag{6}$$

where x_t is the deviation term that might be I(0), and specified as a Gaussian white noise process (with no dependence) or with some type of weak dependence (AR) structure. However, the process x_t may also be generalized to allow for I(1) or even fractional I(d) behaviour. In fact, this is a crucial issue if we want to correctly determine the time trend coefficients, noting that inconsistent estimates are obtained in case of misspecification of the order of integration of the deviation term.

The methodological contribution of the present paper is two-fold: first, we extend the process in (6) to allow for fractional orders of integration in the deviation term. Thus, we examine the significance of the β_1-coefficient (to check for instance if there has been a significant increase in the temperatures), allowing x_t in (6) to be I(d), where d can be 0 (as in the standard case), 1 (a unit root process, I(1)), or any other real value (fractional processes). Moreover, we permit the existence of different time-regimes in the series and thus, we allow for segmented trends, where each of the sub-samples may present different coefficients in (6), along with different orders of integration. Moreover, the times of the breaks will be endogenously determined by the model. Note that fractional integration and structural breaks are issues which are intimately related. Diebold and Inoue (2001), Granger and Hyung (2004) among others showed that structural breaks and I(d) are models which are easily confused.[4] In this article we combine the two approaches in a single framework. Finally, we use a non-parametric approach due to Bloomfield (1973) to describe the time dependence in the (fractionally) d-differenced processes.

The structure of the paper is as follows: In Section 2 we briefly describe the statistical model and the procedure employed for jointly estimating the deterministic terms and the orders of integration for each subsample along with the times of the breaks. Section 3 describes the data. In Section 4 we show the empirical results, while Section 5 concludes the paper.

2. Statistical modelling

We propose in this section an extension of a simple procedure developed by Gil-Alana (2007) for estimating fractional orders of integration with deterministic linear trends and two breaks (and therefore three different regimes) at unknown dates. We assume that y_t is the observed time series, generated by the model

$$y_t = \alpha_1 + \beta_1 t + x_t; \quad (1-L)^{d_1} x_t = u_t, \quad t = 1, ..., T_{b1} \tag{7}$$

$$y_t = \alpha_2 + \beta_2 t + x_t; \quad (1-L)^{d_2} x_t = u_t, \quad t = T_{b1}+1, ..., T_{b2}, \tag{8}$$

$$y_t = \alpha_3 + \beta_3 t + x_t; \quad (1-L)^{d_3} x_t = u_t, \quad t = T_{b2}+1, ..., T, \tag{9}$$

where the α's and the β's are the coefficients corresponding respectively to the intercepts and the linear trends; d_1, d_2 and d_3 may be real values, u_t is I(0) and T_{b1} and T_{b2} are the two times of the breaks that are assumed to be unknown.

This approach is based on the least square principle proposed by Bai and Perron (1998). First, we choose a grid for the values of the fractionally differencing parameters d_1, d_2 and d_3, for example, d_{io} = 0, 0.01, 0.02, …, 1, i = 1, 2, 3. Then, for a given partition $\{T_b\} = \{T_{b1}, T_{b2}\}$ and given d_1, d_2, d_3-values, $(d_{1o}^{(j)}, d_{2o}^{(j)}, d_{3o}^{(j)})$, we estimate the α's and the β's by minimising the sum of squared residuals,

$$\min \sum_{t=1}^{T_{b1}} \left[(1-L)^{d_{1o}^{(j)}} y_t - \alpha_1 \tilde{1}_t(d_{1o}^{(j)}) - \beta_1 \tilde{t}_t(d_{1o}^{(j)}) \right]^2 + \sum_{t=T_{b1}+1}^{T_{b2}} \left[(1-L)^{d_{2o}^{(j)}} y_t - \alpha_2 \tilde{1}_t(d_{2o}^{(j)}) - \beta_2 \tilde{t}_t(d_{2o}^{(j)}) \right]^2$$

$$+ \sum_{t=T_{b2}+1}^{T} \left[(1-L)^{d_{3o}^{(j)}} y_t - \alpha_3 \tilde{1}_t(d_{3o}^{(j)}) - \beta_3 \tilde{t}_t(d_{3o}^{(j)}) \right]^2$$

$$w.r.t. \{\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3\}$$

where $\tilde{1}_t(d_i) = (1-L)^{d_i} 1$, and $\tilde{t}_t(d_i) = (1-L)^{d_i} t$, i = 1, 2, 3.

Let $\hat{\alpha}(T_b; d_{1o}^{(1)}, d_{2o}^{(1)}, d_{3o}^{(1)})$ and $\hat{\beta}(T_b; d_{1o}^{(1)}, d_{2o}^{(1)}, d_{3o}^{(1)})$ denote the resulting estimates for partition $\{T_b\}$ and initial values $d_{1o}^{(1)}, d_{2o}^{(1)}$ and $d_{3o}^{(1)}$. Substituting these estimated values in the objective function, we obtain RSS(T_b; $d_{1o}^{(1)}$, $d_{2o}^{(1)}$, $d_{3o}^{(1)}$), and minimising this expression for all values of d_{1o}, d_{2o} and d_{3o} in the grid we obtain

$$RSS(T_b) \;=\; \arg\min_{\{i,j,k\}}\; RSS(T_b; d_{1o}^{(i)}, d_{2o}^{(j)}, d_{3o}^{(k)}).$$

Then, the estimated break date, $\hat{T}_k = \{\hat{T}_{k1}, \hat{T}_{k2}\}$ is such that

$$\hat{T}_k \;=\; \arg\min_{\,i=1,\dots,m} RSS\,(T_i)\,,$$

where the minimisation is over all partitions T_1, T_2, …, T_m, such that T_i - $T_{i-1} \geq |\varepsilon T|$. The regression parameter estimates are the associated least-squares estimates of the estimated k-partition, i.e.,

$$\hat{\alpha}_i \;=\; \hat{\alpha}_i(\{\hat{T}_k\}),$$

$$\hat{\beta}_i \;=\; \hat{\beta}_i(\{\hat{T}_k\}),$$

and their corresponding differencing parameters,

$$\hat{d}_i \;=\; \hat{d}_i(\{\hat{T}_k\}),$$

for i = 1, 2 and 3.

In Gil-Alana (2007) it is shown that the rates of convergence of the estimates are similar to those in Bai and Perron (1998), since the values are chosen in such a way as to minimise the residual sum of squares and, under the appropriate specification, u_t should follow an I(0) process. Moreover, several Monte Carlo experiments conducted in that study show that the procedure performs well in large samples.

3. The data

The time series analyzed in the paper are the annual land air and sea surface global and (northern and southern) hemispheric temperature anomaly series for the time period

1861-2002, obtained from Jones et al. (2005). These series are continually updated and expanded by P. Jones of the Climatic Research Unite (CRU) with help from colleagues of the CRU and other institutions. Some of the earliest work in producing these temperature series dates back to Jones et al. (1986a, b, c), Jones (1988, 1994), and Jones and Briffa (1992).

[INSERT FIGURE 1 ABOUT HERE]

Figure 1 displays plots of the three time series. It is observed that the three of them present an increasing trend across time. Thermomether records collected in the last 150 years near the surface over land and sea from different parts of the world show an average temperature rising roughly 0.5ºC starting about 150 years ago. The global temperatures show relatively stable temperatures from the beginning of the record through about 1910, with relatively rapid and steady warming through the early 1940s, followed by another period of relatively stable temperatures through the mid-1970s. From this point onward, another rapid rise similar to that in the earlier part of the century is observed. The northern and southern hemisphere annual trend series show some general similarities, e.g., little sign of trends before about 1900, a peak in the early 1940s, and the highest temperatures occurring after 1980. Note that the discontinuous changes observed in these figures have already been examined by many authors. The argument put forward by most climatologists and statisticians is based on the balance between greenhouse gases and sulfur emissions: For much of the initial period, greenhouse gases rose faster than sulfur emissions, which caused temperature to rise. Between the 1940s and 1970s, greenhouse gases and sulfur emissions rose at about the same rate, which caused temperature to remain relatively unchanged. During the final period, legislation, such as the US Clean Air Act slowed the emission of sulfure relative to greenhouse gases, which allowed the warming effect of greenhouse gases to predominate. In the following section we examine if these empirical facts are consistent with the statistical model described in Section 2.

4. The empirical results

In this section we perform the procedure described in Section 2 allowing first for two breaks in the time series data and thus, allowing for three different regimes. Using the global temperatures, the estimated model was:

$$\begin{aligned}
&y_t = \underset{(-6.222)}{-0.46829} + \underset{(2.944)}{0.00266}\,t + x_t; \quad (1-L)^{0.00} x_t = \varepsilon_t, \quad t = 1861, \ldots, T_{b1} = 1871, \\
&y_t = \underset{(-4.566)}{-0.38013} + \underset{(2.338)}{0.00361}\,t + x_t; \quad (1-L)^{0.49} x_t = \varepsilon_t, \quad t = T_{b1}+1, \ldots, T_{b2} = 1974, \\
&y_t = \underset{(-1.281)}{-1.92192} + \underset{(1.954)}{0.01644}\,t + x_t; \quad (1-L)^{0.00} x_t = \varepsilon_t, \quad t = T_{b2}+1, \ldots, T = 2002,
\end{aligned} \tag{10}$$

with the t-values in parenthesis. We observe evidence of fractional integration only over one of the regimes, which is precisely the longest one running from 1872 through 1974. The order of integration is in that case equal to 0.49, implying thus stationarity though close to the nonstationary region ($d \geq 0.5$). For the remaining two subsamples, which are relatively short in the two cases (11 and 28 periods respectively) the orders of integration are exactly 0, implying lack of long memory behaviour. With respect to the time trends, we observe that they are statistically significant in the three cases, and the highest value is the one obtained for the period starting in 1975. We see that, according to this specification, temperatures have increased about 0.26ºC/100 years during the first subsample, at a slightly higher rate (0.36ºC) during the central subsample, and about 1.64ºC/100yrs during the last part of the sample (1975-2002). Note however that the estimates of the time trends for the first and the last subsamples are based on very few observations and therefore they should be not so reliable as those estimates obtained for the second subsample.

We performed the same procedure for the northern and southern anomaly time temperatures, and the resulting models were:

$$y_t = \underset{(-5.530)}{-0.4479} + \underset{(3.349)}{0.00391\,t} + x_t; \quad (1-L)^{0.00} x_t = \varepsilon_t, \quad t = 1861, ..., T_{b1} = 1871,$$

$$y_t = \underset{(-3.790)}{-0.3358} + \underset{(2.107)}{0.00301\,t} + x_t; \quad (1-L)^{0.45} x_t = \varepsilon_t, \quad t = T_{b1}+1, ..., T_{b2} = 1981, \tag{11}$$

$$y_t = \underset{(-9.978)}{-2.8353} + \underset{(10.607)}{0.0237\,t} + x_t; \quad (1-L)^{0.00} x_t = \varepsilon_t, \quad t = T_{b2}+1, ..., T = 2002.$$

for the northern hemispheric temperatures, and

$$y_t = \underset{(-8.055)}{-0.5907} + \underset{(3.901)}{0.0422\,t} + x_t; \quad (1-L)^{0.00} x_t = \varepsilon_t, \quad t = 1861, ..., T_{b1} = 1871,$$

$$y_t = \underset{(-5.936)}{-0.4302} + \underset{(2.489)}{0.00320\,t} + x_t; \quad (1-L)^{0.47} x_t = \varepsilon_t, \quad t = T_{b1}+1, ..., T_{b2} = 1974, \tag{12}$$

$$y_t = \underset{(-9.185)}{-1.5175} + \underset{(9.772)}{0.01300\,t} + x_t; \quad (1-L)^{0.00} x_t = \varepsilon_t, \quad t = T_{b2}+1, ..., T = 2002.$$

for the southern hemisphere. Thus, similarly to the global temperatures, we only found evidence of long memory (fractional integration) for the period located at the central part of the sample, 1872-1981 for the northern temperatures and 1872-1974 for the southern temperatures. The orders of integration are in the two series below 0.5, being slightly higher for the southern temperatures. Once more the slope coefficients are significant in all cases and the highest values correspond to the final part of the sample, 0.0237 and 0.0130 for the northern and southern temperatures respectively. That implies that temperatures have increased 2.37ºC/100yrs in the northern hemisphere and 1.30ºC/100yrs in the southern hemisphere during the final part of the sample.

The procedure implemented so far is based on the assumption that the disturbance term is a white noise process. We also tried with other models allowing for a certain type of time-dependent structure. In particular, we use AR(k) and MA(k) models with k = 1, 2 and 3. In all cases, the break dates were found to take place at the same values as in the previous cases, i.e., 1871 and 1974 for the global and southern temperatures, and 1871 and 1981 for the northern hemisphere. Moreover, since we are mainly interested in the estimation of the time trend coefficients, it is crucial to correctly estimate the orders of integration for each subseries, and here, the estimation results can be very sensitive to the choice of the parameterization of the error term. In what follows, we present the results of the time trend coefficients and the integration orders for each subsample using an alternative way of modelling the I(0) error term. This approach, due to Bloomfield (1973), is non-parametric, and the d-differenced process u_t is exclusively described in terms of its spectral density function, which is given by:

$$f(\sigma^2;\lambda;\tau) \;=\; \frac{\sigma^2}{2\pi}\exp\left(2\sum_{l=0}^{k}\tau_l\cos(\lambda_j\, l)\right). \qquad \textbf{(13)}$$

Like the stationary AR(k) case, this model has exponentially decaying autocorrelations and, using this specification, we do not need to rely on so many parameters as in the ARMA processes, which always results tedious in terms of estimation, testing and model specification.[5]

[INSERT TABLE 1 ABOUT HERE]

Table 1 displays the time trend coefficients and the orders of integration for the second and third subsamples. Note that for the first subsamples, very few observations are available, invalidating thus the analysis based on fractional integration. First, we observe that for the three series, the time trend coefficients are statistically significant at conventional statistical levels.[6] Also, it is observed a substantial increase in the magnitude of the coefficients in the third subsample. For the longest periods (2^{nd} subsamples) the orders of integration are found to be fractional: 0.39 for the global temperatures, and 0.47 for the two hemispheric temperatures. However, for the third subsamples the (95%-) confidence intervals include the I(0) case in the three series.

A drawback of the above specifications is that since we allow for the existence of three different regimes, the procedure determines the breaks in the three series at some extreme values, implying that the first and the final subsamples contain very few observations, which may invalidate the statistical inference about the time trend coefficients. To solve that, we next perform the procedure, assuming that there is a single break on the data. Assuming white noise disturbances, the selected model for the global temperatures is then:

$$y_t = -0.3756 + 0.0035t + x_t; \quad (1-L)^{0.44} x_t = \varepsilon_t, \quad t = 1861, ..., T_b = 1964,$$
$$(-5.634) \quad (2.955)$$
$$(14)$$
$$y_t = -1.9219 + 0.0164t + x_t; \quad (1-L)^{0.00} x_t = \varepsilon_t, \quad t = T_b + 1, ..., T.$$
$$(-12.814) \quad (13.539)$$

Thus, we observe that the break takes place at 1964, and fractional integration (with d = 0.44) only occur for the first subsample. Note that in this case the time trend coefficient is relatively small during the period before 1964 (0.0035), however, it dramatically increases (0.0164) during the period after the break. The selected model for the northern temperatures is

$$y_t = -0.3389 + 0.0030t + x_t; \quad (1-L)^{0.38} x_t = \varepsilon_t, \quad t = 1861, ..., T_b = 1971,$$
$$(-4.811) \quad (2.730)$$
$$(15)$$
$$y_t = -2.835 + 0.0237t + x_t; \quad (1-L)^{0.00} x_t = \varepsilon_t, \quad t = T_b + 1, ..., T.$$
$$(-9.978) \quad (10.607)$$

Thus, for this series, the break happens seven years after than in the previous case of global temperatures. The orders of integration are 0.38 for the first subsample and exactly 0 for the period after the break and, once more, the trend coefficient substantially increase in the second subsample (from 0.30ºC before 1971 to 2.37ºC after that date).

For the southern hemisphere the selected model is:

$$y_t = -0.4743 + 0.0036t + x_t; \quad (1-L)^{0.46} x_t = \varepsilon_t, \quad t = 1861, ..., T_b = 1964,$$
$$(-7.897) \quad (3.350)$$

$$y_t = -1.5175 + 0.0130t + x_t; \quad (1-L)^{0.00} x_t = \varepsilon_t, \quad t = T_b + 1, ..., T,$$
$$(-9.185) \quad (9.772)$$

(16)

which is rather similar to the global case, with the break at 1964, an order of integration for the first subsample slightly smaller than 0.5, and a higher slope coefficient after 1964.

[INSERT TABLE 2 ABOUT HERE]

In the context of a single break, the model of Bloomfield (1973) was also implemented for the error term. The results are displayed in Table 2. We note here that fractional degrees of integration are obtained for the two subseries in the three cases. For the global temperatures the order of integration is similar in the two subsamples, while for the hemipsheric temperatures the values are smaller during the final part of the sample, especially for the southern temperatures. Once more the time trend coefficients are significant in all cases, and there is a substantial increase in the magnitude of the coefficients of the time trends after the break.

Finally, we also performed the procedure for the case of no breaks. Using white noise disturbances, the estimated models were:

$$\underset{}{y_t} = \underset{(-5.494)}{-0.3968} + \underset{(4.632)}{0.0046\,t} + x_t; \quad (1-L)^{0.48} x_t = \varepsilon_t, \quad t = 1, \ldots, T, \tag{17}$$

for the global temperatures.

$$y_t = \underset{(-4.303)}{-0.3710} + \underset{(4.096)}{0.0047\,t} + x_t; \quad (1-L)^{0.46} x_t = \varepsilon_t, \quad t = 1, \ldots, T, \tag{18}$$

for the northern temperatures, and

$$y_t = \underset{(-7.959)}{-0.5024} + \underset{(5.847)}{0.0049\,t} + x_t; \quad (1-L)^{0.46} x_t = \varepsilon_t, \quad t = 1, \ldots, T, \tag{19}$$

for the southern hemisphere anomalies. Here we observe that the order of integration is 0.48 in the global case and 0.46 for the two hemispheres, and the increases in the temperatures are rather similar in the three series, 0.46ºC/100yrs, 0.47 and 0.49 respectively for each case.

[INSERT TABLE 3 ABOUT HERE]

If we implement the model of Bloomfield (1973) for the disturbance term, the results are displayed in Table 3. The fractional differencing parameter d is found to be slightly above 0.5 for the three series implying then nonstationary behaviour, however, the confidence intervals include values ranging from 0.42 to 0.72. The time trend coefficients are again significant, and the values are 0.0046, 0.0048 and 0.0051 respectively for the global, northern and southern temperatures.

5. Concluding comments

In this paper we have examined the global and hemispheric temperatures from a new time series perspective, incorporating fractional orders of integration and with the possibility of three-time regimes for the trends, with the changing times being endogenously determined by the model. Though the results differ depending on the time series used and the model selected (with two, one or no breaks) some common features are obtained in all cases, and they can be summarized as follows: first, we observe that fractional integration occurs in the three anomaly series, with the values for the order of integration being around 0.5 in all cases. This implies that the series are long memory with a strong degree of association across time, and fluctuating around the boundary case between stationary and nonstationary behaviour. Second, the coefficients associated to the time trends are statistically significant in all cases, implying that temperatures have increased across time though at different rates. Thus, the highest increases have taken place during the last twenty or thirty years, increasing around 2.37ºC/100yrs in the northern hemisphere, and about 1.30ºC for the southern and global temperatures. These results are partially consistent with those given by Vinnikov et al. (1990), Jones et al. (1991), Jones and Moberg (2003) and others, which find an increase in the temperatures at a rate of 0.50ºC/100yrs approximately. These papers, however, do not take into account neither fractional integration nor segmented trends.

Finally, the results about the orders of integration of the series obtained in the paper solve partly the controversy about the stationary/nonstationary nature of the series. Thus, while some authors assume that global and hemispheric temperatures are stationary I(0), others claim that they are I(1). Thus, for example, Stern and Kaufmann (2000) and Kaufmann and Stern (2002) find that these series are I(1), which permit them to make statistical inference based on cointegration. However, in that study, they only obtain evidence of unit roots in the temperatures when using the Kwiatkowski et al. (KPSS, 1992) procedure, while using other methods like Phillips and Perron (PP, 1988) and Schmidt and Phillips (SP, 1992), evidence supports the I(0) trend stationary representation. Note, however, that the procedures employed in that paper only distinguish between I(0) and I(1) representation and do not take into account the possibility of fractional integration.[7] Our results clearly support the view that the three series are I(d) with d about 0.5, and also show a clear increase in the temperatures in the last twenty or thirty years. Various theories have been offered to explain the apparent rise in temperature in recent decades. This paper presents evidence that there has in fact been an increase, but makes no claim as to the cause.

Endnotes

1. A process $\{x_t\}$ is said to be (covariance) stationary (or second order stationary) if the mean and the variance do not depend on time and the covariance between any two observations depends on the distance between them but not on their specific location in time. This is a minimal requirement for statistical inference in time series.

2. Examples of I(0) processes are the white noise case, the stationary AR, Moving Average (MA), stationary ARMA, etc.

3. Some of these authors argue that the presence of unit roots in temperature is caused by the I(1) trends observed in radiative forcing. Using multivariate techniques, it is argued that temperature cointegrates with many components of radiative forcing. However, these techniques concentrate exclusively on the cases of stationary I(0) and nonstationary I(1) and do not take into account other plausible fractional alternatives.

4. In fact, the existence of breaks may lead to spurious findings of long memory. Lobato and Savin (1998) argue that structural breaks may be responsible for the long memory in return volatility processes, and Engle and Smith (1999) investigated the relationship between structural breaks and long memory using a simple model where the data generating process consists of a mean process and a stationary error.

5. See Gil-Alana (2004) for an explanation of the exponential spectral model of Bloomfield (1973) in the context of fractional integration.

6. We use Newey-West HAC corrections for the estimation of the variance-covariance matrix of the regression estimates.

7. Diebold and Rudebusch (1991), Hassler and Wolters (1994) and others have showed that the classic methods for testing unit roots (e.g., Dickey and Fuller, ADF,

1979; Phillips and Perron, 1988; Kwiatkowski et al., KPSS, 1992) are seriously biased if the alternatives are of a fractional form.

Acknowledgements

Comments of three anonymous referees are gratefully acknowledged. The author also acknowledges financial support from the Ministerio de Ciencia y Tecnologia (SEJ2005-07657, Spain).

References

Bai, J. and P. Perron, 1998, Estimating and testing linear models with multiple structural changes, Econometrica 66: 47-78.

Baillie, R.T., 1996, Long Memory Processes and Fractional Integration in Econometrics, Journal of Econometrics 73: 5-59.

Beran, J., 1994, Statistics for Long-Memory Processes. London: Chapman and Hall.

Bloomfield, P., 1973, An exponential model for the spectrum of a scalar time series, Biometrika 60: 217-226.

Bloomfield, P., 1992, Trends in global temperatures, Climate Change 21: 275-287.

Dickey, D. A. and W. A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association 74: 427-431.

Diebold, F.X. and A. Inoue, 2001, Long memory and regime switching, Journal of Econometrics 105: 131-159.

Diebold, F.S. and G.D. Rudebusch, 1991, On the power of Dickey-Fuller tests against fractional alternatives, Economic Letters 35: 155-160.

Doukhan, P., G. Oppenheim and M.S. Taqqu, 2003, Theory and applications of long range depedence, Birkhäuser, Basel.

Dueker, M.J. and P.K. Asea, 1998, Non-monotonic long memory dynamics in black market premia, Typescript, The Federal Reserve Bank of St. Louis.

Engle, R.F. and Smith, A.D., 1999, Stochastic permanent breaks, Review of Economics and Statistics 81: 553-574.

Gil-Alana, L.A., 2003, An application of fractional integration to a long temperature time series, International Journal of Climatology 23: 1699-1710.

Gil-Alana, L.A., 2004, The use of Bloomfield (1973) model as an approximation to ARMA processes in the context of fractional integration, Mathematical and Computer Modelling 39: 429-436.

Gil-Alana, L.A., 2005, Statistical model of the temperatures in the northern hemisphere using fractional integration techniques, Journal of Climate 18: 5357-5369.

Gil-Alana, L.A., 2007, Fractional integration and structural breaks at unknown periods of time, forthcoming in Journal of Time Series Analysis.

Granger, Clive W.J. and Z. Ding, 1996, Varieties of long memory models, Journal of Econometrics 73: 61-77.

Granger, C.W.J. and Hyung, N., 2004, Occasional structural break and long memory with an application to the S&P 500 absolute stock returns, Journal of Empirical Finance 11: 399-421.

Hassler, U. and J.Wolters, 1994, On the power of unit root tests against fractional alternatives, Economic Letters 45: 1-5.

Jones, P.D. 1988. Hemispheric surface air temperature variations: Recent trends and an update to 1987. Journal of Climate 1: 654-60.

Jones, P. D. 1994. Hemispheric surface air temperature variations: A reanalysis and an update to 1993. Journal of Climate 7(11): 1794-1802.

Jones, P. and T.M.L. Wigley. "Global and Hemispheric Anomalies." In Trends '91: A Compendium of Data on Global Change, edited by T.A. Boden, R.J. Sepanski, and F.W. Stoss. Oak Ridge: Oak Ridge National Laboratory, 512-517.

Jones, P.D., and K.R. Briffa. 1992. Global surface air temperature variations during the twentieth century: Part 1, spatial, temporal and seasonal details. The Holocene 2: 165-79.

Jones, P.D. and A. Moberg, 2003, Hemispheric and large scale surface air temperature variations. An extended revision and update to 2001, Journal of Climate 16: 206-223.

Jones, P.D., S.C.B. Raper, R.S. Bradley, H.F. Diaz, P.M. Kelly, and T.M.L. Wigley. 1986a. Northern Hemisphere surface air temperature variations: 1851-1984. Journal of Climate and Applied Meteorology 25(2): 161-179.

Jones, P.D., S.C.B. Raper, and T.M.L. Wigley. 1986b. Southern Hemisphere surface air temperature variations: 1851-1984. Journal of Climate and Applied Meteorology 25(9): 1213-1230.

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

Jones, P.D., D.E. Parker, T.J. Osborn and K.R. Briffa, 2005, Global and hemispheric temperature anomalies_land and marine instrumental records, In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information. Analysis Center, Oak Ridge, National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn, U.S.A.

Kaufmann, R.K., H. Kauppi and J.H. Stock, 2006, The relation between radiative forcing and temperature: what do statistical analyses of the observational record measure? Climate Change 77: 279-289.

Kaufmann, R.K. and D.I. Stern, 2002, Cointegration analysis of hemispheric temperature relationsh, Journal of Geophysical Research 107.

Koscielny-Bunde, E., A. Bunde, S. Havlin, H.E. Roman, Y. Goldreich and H.J. Schellnhuber, 1998, Indication of a universal persistence law governing atmospheric variability, Physics Review Letters 81: 729-732.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin, 1992, Testing the null hypothesis of stationarity against the alternative of a unit root, Journal of Econometrics 54: 159-178.

Lewis, P.A.W. and B.K. Ray, 1997, Modelling long-range dependence, nonlinearity and periodic phenomena in sea surface temperatures using TSMARS, Journal of the American Statistical Association 92: 881-893.

Lobato, I.N. and Savin, N.E., 1998, Real and spurious long memory properties of stock market data, Journal of Business and Economic Statistics 16: 261-268.

Maraun, D., H.W. Rust and H. Timmer, 2004, Tempting long memory on the interpretation of DFA results, Nonlinear Processes in Geophysics 11: 495-503.

Newey, W.K. and West, K.D., 1987, A simple, positive definite, heteroscedasticity and autocorrelation consistent covariance matrix, Econometrica 55: 703-708.

Pelletier, J. and D. Turcotte, 1999, Self-affine time series II. Aplications and Models, Advances in Geophysics 40: 91-166.

Pethkar, J.S. and A.M. Selvam, 1997, Nonlinear dynamics and chaos. Applications for prediction of weather and climate, Proc. TROPMET 97, Bangalare, India.

Percival, D.B., J.E. Overland and H.O. Mofjeld, 2004, Modelling North Pacific climate time series, to appear in Time Series Analysis and Applications to Geophysical Systems, edited by D.R. Brillinger, E.A. Robinson and F.P. Schoenberg, Springer-Verlag.

Phillips, P.C.B. and P. Perron, 1988, Testing for a unit root in a time series regression, Biometrika 75: 335-346.

Robinson, P.M., 1994, Time Series with Strong Dependence, Invited Paper 1990 World Congress of the Econometric Society, Advances in Econometrics: Sixth World Congress, Volume 1 (C.A. Sims, ed.) Cambridge University Press, 47-96.

Robinson, P.M., 2003, Long memory time series, Time Series with Long Memory, (P.M. Robinson, ed.), Oxford University Press, Oxford, 1-48.

Schmidt, P. and P.C.B. Phillips, 1992, LM test for a unit root in the presence of deterministic trends, Oxford Bulletin of Economics and Statistics 54: 257-287.

Smith, R.L., 1993, Long-range dependence and global warming, Statistics for the Environment (ed. Barnett, V. And Turkman, K.F., pub. J. Wiley), 141-161.

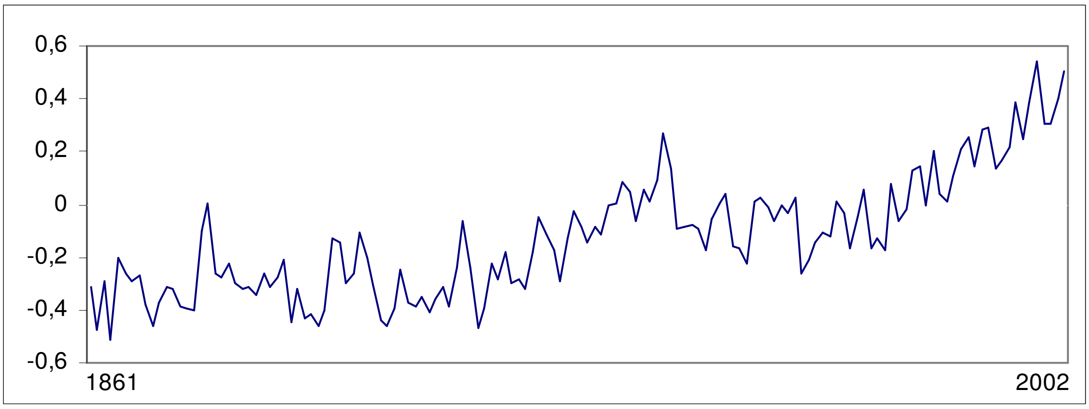
Stern, I.D. and R.K. Kaufmann, 2000, Is there a global warming signal in hemispheric temperature series? Climatic Change 47: 411-438.

Vinnikov, K.Ya, P. Ya Groisman and K.M. Lugina, 1990, Empirical data on contemporary global climate changes (temperature and precipitation), Journal of Climate 3: 662-677.

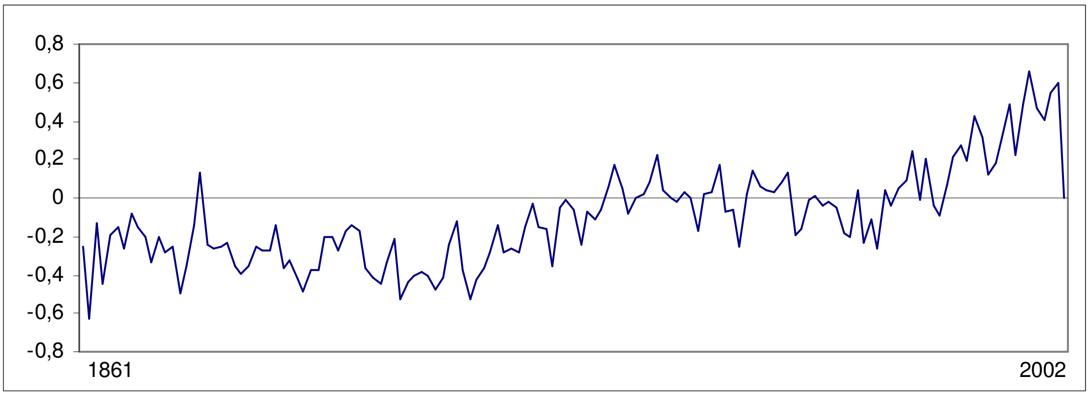
Woodward, W.A. and H.L. Gray, 1995, Selecting a model for detecting the presence of a trend, Journal of Climate 8: 1929-1937.

FIGURE 1

Annual land air and sea surface temperature anomalies: Globe



Annual land air and sea surface temperature anomalies: Northern Hemisphere



Annual land air and sea surface temperature anomalies: Southern Hemisphere

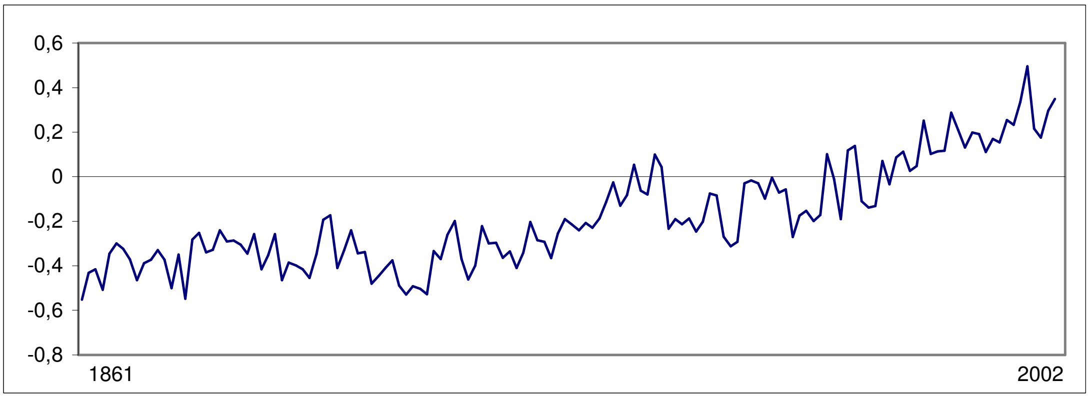


Table 1: Estimates of d and the time trends using the Bloomfield approach

2-breaks	2nd sub-sample		3rd sub-sample	
	Time trend	Integ. Order	Time trend	Integ. Order
Global temp.	0.0029 (2.885)	0.39 [0.26, 0.55]	0.0189 (9.953)	-0.02 [-0.29, 0.46]
Northern temp.	0.0031 (2.176)	0.47 [0.36, 0.61]	0.0283 (6.988)	-0.01 [-0.26, 0.35]
Southern temp.	0.0032 (2.607)	0.47 [0.32, 0.68]	0.0130 (8.671)	-0.14 [-0.62, 0.29]

The values in parenthesis refer to the t-values based on the HAC correction procedure of Newey and West (1987). The values in brackets are the 95% confidence bands for the fractional differencing parameters.

Table 2: Estimates of d and the time trends using the Bloomfield approach

1-break	1nd sub-sample		2rd sub-sample	
	Time trend	Integ. Order	Time trend	Integ. Order
Global temp.	0.0030 (2.452)	0.45 [0.31, 0.66]	0.0166 (4.981)	0.46 [0.13, 0.71]
Northern temp.	0.0026 (1.684)	0.49 [0.37, 0.65]	0.0242 (4.650)	0.41 [0.08, 0.73]
Southern temp.	0.0033 (2.558)	0.51 [0.34, 0.79]	0.0127 (6.467)	0.17 [-0.13, 0.45]

Table 3: Estimates of d and the time trends using the Bloomfield approach

No breaks	Time trend coefficients	Orders of integration
Global temperatures	0.0046 (4.048)	0.52 [0.42, 0.64]
Northern hemisph. temp.	0.0048 (3.148)	0.54 [0.44, 0.67]
Southern hemisph. temp.	0.0051 (3.765)	0.59 [0.49, 0.72]