The Search for Warming in Global Temperatures: Data, Methods and Unresolved Questions

Ross McKitrick
University of Guelph
Department of Economics
August, 2001

rmckitri@uoguelph.ca

Abstract

Global climate change policy is motivated by recent warnings that the world is heating up by a rate of at least 0.14 degrees C/decade. Global temperatures measured at the surface contain a positive trend of about 0.04 ^{o}C/decade since 1856. Evidence for an acceleration of warming after1979 appears in the surface thermometer network, but not in data collected by weather satellites, weather balloons or indirect temperature proxies like tree rings and ice cores. The conflict among these data sets is at the core of ongoing debates about whether "greenhouse gases" are really causing global warming. Different methods for computing global temperatures are compared, and techniques for estimating and interpreting trends are discussed. It is argued here that costly policy commitments like Kyoto should be delayed at least until the discrepancies among the different data sets are resolved.

Acknowledgments: I thank, without implicating, John Galbraith, S. Fred Singer, Olavi Kärner and Christopher Essex for helpful discussions. Funding from the University of Guelph URA program is gratefully acknowledged.

Key words: Global warming, trend estimation

JEL Codes: Q2

The Search for Warming in Global Temperatures: Data, Methods and Unresolved Questions

> "Globally, average surface-air temperatures are projected to rise 2.5 F to 10.4 F between 1990 and 2100, with most land areas warming more than the global average by up to 40 percent. A simple climate model simulating the response of seven complex global-climate models projects globally averaged surface-temperature increases of 2.5 F to 10.4 F between 1990 and 2100 for the full range of plausible trends in greenhouse gas and sulfur-dioxide emissions."
>
> (Dr. Robert Watson, Chair, Intergovernmental Panel on Climate Change, 2001)

1. Introduction

The stylized facts of global warming are well known. Carbon dioxide emissions have been rising over the twentieth century (Marland *et. al* 1999). The concentration of CO_2 in the atmosphere has risen, from about 285 parts per million in 1850 to 369 parts per million in 2000 (Keeling and Whorff 2000b). Computer simulations of the climate system predict that this should be causing the atmosphere to warm, in turn heating up the earth's surface. This paper takes current range of projected warming as a testable hypothesis and looks at whether the data support it. A complication is that different types of data for measuring warming are available, but while some show a warming trend others do not. Also, different techniques for estimating trends are available, and the choice of statistical model has nontrivial implications for interpreting the evolution of climate. Both of these points are essential for evaluating claims and counter-claims about the magnitude of the global warming problem, but they have been widely overlooked outside specialist literature. Their considerable implications for the policy debate motivates this review.

The discussion begins by reviewing strengths and weaknesses of the major sources of data on global temperatures, then discusses some issues involved in measuring trends. Updated global surface and atmospheric temperature anomaly[1] data are then examined and linear trends are computed over the past

---

[1] Temperatures are reported as "anomalies" (deviations from a local mean) rather than levels. At each recording station the mean calculated over a standard interval is removed and the residual is the station "anomaly". These are compiled into global averages using weights to control for grid size and latitude.

150 years and the past 20 years. The temperature trend for the past century is about 42 standard deviations below the central value of the forecast warming rates for the present century. There is mixed evidence for a recent acceleration. Data from the surface thermometer network suggests that since 1979 the warming trend has risen to about the minimum of the range projected by climate models, but data collected by satellites show no such change. Additional data from weather balloons and indirect temperature proxies are examined to see if they can resolve the discrepancy. Neither data set confirms an acceleration of lower atmospheric warming. Taken together the evidence suggests that global temperatures are rising, but the acceleration projected by the IPCC cannot currently be validated. The world seems to be warming at a rate of under half a degree (Celsius) per century.

2. Isolating a Hypothesis for Testing

An important point to note from the outset is that detecting a warming trend in 20th century temperature data does not resolve the question of cause. After all, the world warms and cools over time due to natural processes (think, for instance, of the transition out of the last ice age). Of interest to us is whether there is a warming trend in the 20th century that can be distinguished from natural variation during the period for which we have global temperature data of an acceptable quality, which might be attributable to human activity. Statistical significance of a slope coefficient is neither necessary nor sufficient to identify an anthropogenic influence. The relevant hypothesis is not simply whether there is warming, but whether there is as much warming as climate models tell us to expect.

Twentieth century data are complicated by the presence of at least four natural effects which contributed to climate warming. First, the sun brightened over the past few centuries and reached a 400-year peak irradiance in the mid-1940s, then again in the late 1970s and 1980s (Hoyt and Schatten 1993). While the exact climate implications remain much in dispute it is widely acknowledged that at least some of

---

this century's warming is due to this change in solar output, which accounts for part of the recovery from the cold centuries of the "Little Ice Age" (1300-1850). Second, an ocean mixing cycle with an 1800-year period induced by variations in the lunar orbit is in a weakening phase which began 500 years ago, and will produce a global warming trend for another 400 years before reversing into a cooling mode (Keeling and Whorff 2000a). This too accounts for an unspecified amount of warming in the 20th century temperature record. Third, there are other low-frequency climate cycles of approximately 1500-year duration which have recently been identified as being in a positive phase. These include the Dansgaard-Oeschger cycle, the Bond-Campbell cycle and the Heinrich-Bond cycle (Yu and Wright 2001, Campbell *et. al.* 1998). At present it is not certain that these cycles are distinct, but the existence of at least one is now well-established. An implication is that such cycles will show up in short time-series like the 20th century as an upward trend. Finally, some Northern Hemisphere warming since 1980 has been attributed to the intensification of a climate variation known as the Arctic Oscillation (Thompson *et. al.* 2000), a change which is widely (but not universally) viewed as having natural causes. These all add to warming in twentieth-century temperature data, but the magnitude of their contributions is not known and therefore cannot be subtracted out. Consequently, mere identification of an upward trend in 20th century anomalies does not suffice to show that humans are changing the climate.

Attempts to isolate the role of greenhouse gases rely on computer simulations of the past and future climate. The literature associated with this field of endeavour is, of course, vast, but a few key numbers can be singled out. A decade ago, the First Assessment Report of the Intergovernmental Panel on Climate Change (IPCC 1990) concluded the climate would warm by between 1.5 and 4.5 °C in response to a doubling of carbon dioxide concentrations, which it predicted would happen over the next century. This hypothesized range of sensitivities has not changed in subsequent reports. In the recent Third Assessment Report or TAR (IPCC 2001) the projected temperature change for the post-1990 period is 1.4—5.8 °C per century. Compared to the range of temperature projections (1—3.5 °C) in the Second Assessment Report

(IPCC 1995) both the low end and the high end increased, the latter rising by over 2 degrees. The increase of the top end was primarily due to consideration of a new fuel-intensive emissions scenario, called A1FI.

There are semantic subtleties in IPCC documents on how model outputs should be interpreted. The terms “forecast” and “prediction” are specifically eschewed, instead the simulations are referred to as “projections” or “illustrative scenarios”. But the semantics appear have been lost on policy-makers. Shortly after the release of the TAR Summaries, UN Secretary-General Koffi Annan traveled to Dhaka Bangladesh to speak about global warming, offering the following warnings (emphasis added):

> The United Nations Intergovernmental Panel on Climate Change -- which includes some brilliant scientists from Bangladesh -- has just released its latest *forecast*. The Panel's portrait of a warming world makes for chilling reading. It warns of adverse consequences such as the melting of glaciers and polar ice caps, leading to rising sea levels. It *predicts* more extreme droughts, floods and storms, and significant changes in the functioning of critical ecological systems such as coral reefs and forests. Warmer and wetter conditions would increase the spread of infectious diseases such as malaria and yellow fever. And the inundation of low-lying islands and coastal areas could lead to the displacement of hundreds of millions of people…. The climate Panel's report says that sea-level rise could cause the disappearance of vast swaths of this region, and along with them species such as the famed Bengal tiger.

(Annan 2001).

If the Secretary-General of the UN treats the Panel’s report as forecasts and predictions, it is presumably fair for researchers to at least treat them as testable hypotheses. The question then becomes whether the warming process is *currently* visible. The response to greenhouse gas forcing in a climate model is relatively rapid,[2] and it is not the case that climate simulations predict the warming will only begin later this century. It is now 13 years since IPCC author and NASA climatologist Dr. James Hansen told the US Senate Committee on Science and Technology that he could state with "99 percent confidence" that "The greenhouse effect has been detected and is changing our climate now." (quoted in Revkin, 1988). Hengeveld (2000), in a recent report to the Canadian government, uses a well-known climate model to predict that global temperatures will increase 4.5 ^{o}C starting in 1985 and continuing past 2100. A warming rate of 0.2 ^{o}C per decade is expected from 1985 to 2010, rising thereafter. The quotation from IPCC Chairman Dr. Robert Watson at the start of this paper clearly implies that anthropogenic warming has been underway for a decade, and should therefore be observable now.

So for the present purposes, the IPCC-projected warming range will be treated as a testable hypothesis, since policy makers have acted on the assumption that it is a prediction, not merely an "illustrative scenario". Also, we shall take IPCC pronouncements to mean that a global trend of at least 0.14 ^{o}C per decade should be observable in the more recent data. Identification of a warming trend of this magnitude does not settle the question of cause, since there are natural changes afoot, but it does provide a specific target for which to test. We now consider the available data.

3. Measurement issues

3.1 The Surface Record

Surface temperatures are measured on land by weather stations and at sea by buoys and merchant vessels. The monitored portion of the earth's surface is called the "temperature field" (Kärner 2001). The longest published records of global temperatures begins in the early 1850s. The most widely-used series spans 1856 to the present, and is produced by the Climate Research Unit (CRU) of the University of East Anglia. Putting together such a series involves careful station-by-station scrutiny, reviewing both the data

---

[2] The Transient Climate Response reported in IPCC (2001) is between 50 and 87 percent. It measures the amount of warming contemporaneous with the increase in greenhouse gases.

and the written documentation about the location and operation of each station. The physical credibility of long historical samples is impaired by the fact that the temperature field has changed continuously over the past 150 years. As of 1860, only 18 percent of the surface of the globe had thermometer coverage, and by 1900 this had only risen to 40 percent (Kärner 2001). Another major data collection which shares records with the CRU is called the Global Historical Climatology Network (GHCN). Peterson and Vose (1997) show that in the GHCN, most of the temperature measurement stations that operated continuously over the 20th century are concentrated in the US and Europe. Of particular concern, oceans, which cover three-quarters of the earth's surface, yield very little air temperature data. The Southern Hemisphere as a whole is poorly covered prior to the late 1950's. As well, the polar regions, Greenland, Siberia, the northern half of Africa, the Arabian peninsula, the interior of Australia and a large swath of central Asia down into India are subject to very thin sampling. The temperature field covered by the Global Historical Climatology Network attained its largest extent in the early 1970s, at which time about 6000 land-based weather stations were in operation. Since then, the field has suffered dramatic decay, especially in the former Soviet Union, and today only about 2000 weather stations are in operation globally, covering only about 50 percent of the earth's surface (Peterson and Vose 1997, Figure 2).

A concern expressed by critics of the surface data is that the stations closed over the past 25 years were disproportionately in remote and underdeveloped areas, and about 90 percent of the land-based data now being used to construct global averages are sampled in cities. This raises the possibility of contamination due to "urban heat island" effects. In a pure greenspace, only a fraction of incoming solar energy is converted into sensible heat, while the rest is used in photosynthesis to produce plant biomass and some evaporates surface moisture to produce water vapour (Lockwood 1979). But when roads and structures replace greenery and cover the earth with dry surfaces, all the solar energy is converted to sensible heat, raising air temperatures. Buildings and traffic also generate and store heat all day, and together with black surfaces release it at night, creating strong updrafts of heated air. Heat island effects can certainly be felt in urban areas, but can even be noticed in rural areas if there are enough roads and buildings in the vicinity of the weather station. In Europe, over half the temperature stations available as of 1850 are in towns that now have at least 500,000 inhabitants (Böhm 1998). Globally, airports are common measurement sites, and there are many cases where an airport was located in a rural or semi-rural location as of the end of WWII, but is now urbanized.

There is no generally accepted way to measure urban heat island bias at individual temperature stations. One technique is to compare reliable measurements in adjacent urban and greenspace areas over a long period. Owing to the limited number of locations where such comparisons are possible, much of the research on urban effects has focused on the US and Europe. Studies that have used this approach to remove estimated urban heat island effects from US and European data show the 1930's and early 1940's to be the warmest years of the twentieth century (Singer 1999). Similar results have been found for South Africa (Balling and Hughes 1996), Alaska (Magee, Curtis and Wendler 1999) and elsewhere.

Another approach uses a regression method to predict the bias as a function of population growth (Peterson and Vose 1997). The adequacy of this approach has been challenged. In a study of temperature records in post-war Vienna, Böhm (1998) finds the overall urban temperature bias grew significantly between 1951 and 1995, adding between 0.1 and 0.6 ^{o}C to measured temperatures depending on location. Yet the population of Vienna fell over the period. The heat bias arose due to increased energy use, increased traffic, and the fact that forest and grassland cover in the metropolitan area decreased from 63 percent to 48 percent.

In comparison to the problems of getting reliable measurements over land, getting air temperatures over the oceans is even harder. Since oceans cover three-quarters of the globe[3] a credible instrumental series must have accurate readings of marine air temperatures. However, the area is vast and the number of fixtures (like buoys) which could monitor air temperatures is very limited. Therefore sea surface

---

[3] In addition, large inland bodies of water like the Great Lakes cover significant areas but have very few weather stations on their surfaces.

temperature (SST) records, i.e. water temperatures, taken by merchant marine vessels are used as proxies for marine air temperature.

Regular coverage of water temperatures is limited to shipping lanes, and measurement techniques have not been uniform over the century. At first, sailing ships dipped a bucket into the ocean and used a glass thermometer to take the temperature reading. There were several changes in thermometer and bucket design, as well as uneven practices on whether water was taken from the side of the ship or in the wake behind it, and how long the thermometer was left in the bucket before taking the reading. After the advent of steam power in the early 20th century, ships began automatic measurement of the temperature of water injected into the engine cooling system. This water comes from below the surface, unlike bucket-sampled water, and depending on where in the ship's interior the temperature is read, the water may be warmed by the ship and/or the engine. Over much of the 20th century, the two methods have been used together, but in proportions which were never recorded. The proportion of engine-water readings has grown, adding an artificial warming to the ocean data which some marine scientists have argued may be as much as 0.5 degrees C over the century (Monastersky 1989).

In addition there are new concerns that water temperatures do not follow the same trends as air temperatures, even at the same location. Christy *et. al.* (2001) examine post-1979 SST and Marine Air Temperature (MAT) data from ships, as well as data from weather satellites, weather balloons and a network of buoys in the tropical Pacific. The buoy network data are especially useful since they measure temperatures at one meter below the surface and three meters above it in the same location. In all comparisons of SST and air temperature it is found that the ocean has been warming relative to the air. Moreover, three of the four temperature data sets (satellite, balloon and buoys) indicate marine air temperatures just above the ocean surface have been *cooling* throughout the tropics at an average rate of between 0.01 and 0.06 ^{o}C per decade since 1979, even while the SST data has shown warming. Yet it is the SST data which are used by the CRU to compile global temperatures. The authors re-calculated global average temperatures over intervals where air temperature data were available instead of SST, and the global trend was reduced by 0.05 ^{o}C per decade.

The CRU data are the most widely-used global surface temperature estimates and are made available on-line by the UK Hadley Centre. Monthly global surface temperature anomalies from 1856-2000 are shown in Figure 2. Following the end of the Little Ice Age (1350-1850) the climate experienced a rapid, though uneven increase in average temperatures of about 0.4 ^{o}C, lasting until about 1940. Temperatures cooled slightly for the next four decades and then rose about 0.4 ^{o}C in the last two decades.

3.2 Satellite Data

The lower portion of the atmosphere, up to 10 km at the poles and 16 km in the tropics, is called the troposphere. Climate models project warming to be as strong or stronger here than on the surface (NRC 2000) since this is the layer where greenhouse gases mix and change the optics of the air. Since 1979, National Oceanic and Atmospheric Administration (NOAA) satellites have used microwave sounding units (MSU) to measure the intensity of microwave emissions from oxygen molecules in the air. The MSUs take over 30,000 readings per day, and each reading yields the average “brightness” of a 75,000 cubic kilometer sample of air. Since microwave emissions vary with temperature according to known physical relationships, the average temperature level for each air sample can be computed. This technique provides daily measurements with a continuous temperature field throughout the lower troposphere covering 95% of the earth.[4] In contrast to the surface data, satellite measurements have shown no significant warming in the troposphere since 1979. Prior to the strong 1998 El Nino event there was a small and insignificant cooling trend. Figure 3 shows monthly temperature anomalies as measured by weather satellites from January 1979 to January 2001.

The discrepancy between surface and tropospheric trends has attracted considerable attention. Much of the debate about whether the world is warming comes down to the question of which of these data sets should be given more prominence. Technical challenges associated with obtaining reliable temperature measures from the satellite include orbital decay and east-west shift, heating of the instrument surface and merging series when one satellite is removed from service and another replaces it (Hurrell *et. al.* 2000, Christy, *et. al.* 1997). Currently available series reflect corrections for these effects, as best as can be done. The corrections have involved changes up and down on a magnitude of a few hundredths of a degree per decade. Another limitation is that the readings are only available for the past twenty-one years, which is generally considered too short a span to establish a climatological pattern. The advantages of the satellite data are that they provide near-global coverage over a consistent temperature field, their measurements are not contaminated by urban waste heat, and they measure temperatures in the troposphere itself, where the mixing of greenhouse gases occurs and where temperature changes should be apparent.

2.3 Balloon Data

The US National Oceanic and Atmospheric Administration (NOAA) maintains a set of quarterly temperature records taken by an international network of 63 weather balloons (Angell 2000). The balloons carry shoebox-sized containers called *radiosondes*, a term coined to describe a platform of meteorological instruments which transmit data to a receiver on the ground. The weather balloons ascend from the surface to the mid-stratosphere and transmit, among other things, temperature data by altitude, as measured by an electronic device called a thermistor. The surface coverage is not global, instead it roughly follows the distribution of the surface data. Europe, southern Canada and the US are covered, but Mexico and South America, Africa, Central Asia, Greenland and the polar regions have relatively few observations, and those that are available tend to come from cities in these places. The temperature anomalies from the low troposphere for 1979-2000 are graphed in Figure 4. These are averages over the 850-300 millibar layer, extending approximately 1-10 km up into the atmosphere.

---

[4] See http://vortex.nsstc.uah.edu/essc/msu/background.html for an explanation of satellite measurements.

3.4 Temperature Proxy Data

Information on surface temperatures also comes from indirect sources, such as tree ring growth rates and the thickness of ice layers in glaciers. While somewhat noisy, these proxy measures do, under certain circumstances, reveal local temperature changes on annual scales. There is a great deal of interest in these measures because they can potentially provide long-term historical temperature reconstructions against which to compare the 20th century record. Of additional interest for the purpose of this study is that many of the proxy records come from areas remote from human habitation, and are therefore unaffected by urban waste heat, and depending on how recent the sample is it can reveal information about local postwar temperature trends. After discussing the empirical results based on the other three data types some of the findings from proxy reconstructions will be surveyed.

4. Trend estimation

Statistical models allow us to use information in a limited sample to infer something about the larger population from which the sample was drawn. In the case of geophysical time-series data, unique challenges arise because sample length is extremely small in comparison to the time scale of the physical system. Having 150 years of temperature data (ignoring for the moment quality problems in the early decades) does not suffice to reveal the character of processes that might play themselves out over thousands or hundreds of thousands of years. This raises some problems of interpretation that are the focus of continuing interest among statisticians who use climate data.

To simplify somewhat, there are four simple time-series models that can be used, alone or in combination, to relate observations on a given period to "time" as a whole. The simplest is the *random walk* model

$$A(t) = A(t-1) + e(t) \tag{1}$$

where $A(t)$ is the value of the variable being studied (e.g. a temperature anomaly) at time t, and $e(t)$ is a random number with a mean of zero representing natural, unpredictable shocks to A between observation periods. Equation (1) states that each period's value is just last period's plus a random, unpredictable amount. Another time series equation, similar to (1), is called an *autoregressive* model:

$$A(t) = q\ A(t-1) + e(t) \qquad (2)$$

where q is a parameter between –1 and +1. (2) implies that the influence of random changes dies out over time, whereas in (1) a change in $A(t)$ is permanent. That is, if $e(t)$ were zero for every period after t, in (2) $A(t)$ would taper off to zero whereas in (1) it would remain at $A(t)$. Since there is only one lagged value in (2) it is referred to as a *first-order* autoregressive process, or AR(1) for short. Additional lagged terms can be added in if needed.

A third process is called a *moving average* and a first-order version is written

$$A(t) = ge(t-1) + e(t) \qquad (3).$$

This process implies that the value of the variable each period is the sum of this period's random shock and last period's random shock, the latter weighted by a parameter g. As with (2), if random shocks did not occur in the future this process would dampen down to zero over time, as long as g is between –1 and +1. The fourth simple time-series process is the trend model

$$A(t) = b\ t + e(t) \qquad (4).$$

This one states that $A(t)$ grows each period by the amount b plus a random shock.

The above are all written in their simplest forms, but modifications can include adding constant terms and, for (2) and (3), extending the number of lagged terms on the right hand side. What is challenging for empirical work is that many time series could be plausibly described by any of the models (1)-(4), but the interpretations differ substantially. (1), for instance, suggests that the system represented by $A(t)$ moves around at random, and past changes in the system give us no information about future changes. (4), on the other hand, suggests that the system follows a steady, predictable pattern as time passes, and past changes can be used to predict future values. In the case of (2) and (3) both models imply that whatever the current value of A the system "wants" to return to a zero state, and if it were not for random fluctuations it would eventually do so.

If global average temperatures follow a process like (1), it amounts to a very different thing than if they follow a process like (2) or (3) or (4). Data generated by a random walk can sometimes seem to have trends over long periods, even though future temperature changes are formally unpredictable. One way modelers decide among (1)-(4) is to combine them all into a framework called *ARIMAX* (which stands for Autoregressive Integrated Moving Average With Trend) and then gradually reduce the dimensions of the system until test statistics reject any further simplifications. The resulting models might, for instance, combine (2) with p lagged terms, (3) with q lagged terms and a trend, in which it would be called an ARMAX(p,q) model.

Woodward and Gray (1995) and Kärner (1996) examine global temperature data and conclude that they follow a random walk, implying that any "trends" in the past teach us nothing about the future evolution of climate. Zheng and Basher (1999) review studies published up to 1997 and find a general conclusion that the long-run trend in surface temperatures is not statistically significant. Noting that many have used a simple trend model (4) rather than an ARIMA framework they remark:

> "If such models [i.e. ARIMA] were indeed correct, then the fitting of a model comprising a trend and stationary residuals could result in the erroneous detection of a trend. These authors are cautious in their conclusions, none claiming that global warming does not exist, but they argue that statistically we cannot rule out the possibility that the past century's rise in global temperatures is simply the result of natural variability and thus may not continue into the future. Their results present a disconcerting challenge to the conventional view that global temperatures are rising because of some external forcing, probably associated with anthropogenic factors, and are likely to continue to rise." (p. 2347)

Zheng and Basher then use an ARIMA model with some additional explanatory variables (sunspots, volcanic dust, etc.) and find a significant global trend of 0.046 ^{o}C/decade on data spanning 1861-1993.

Other authors have used so-called *unit root* tests, developed for financial market analysis, to look for evidence of random walk behaviour. Stern and Kaufman (1997) find two classes of tests (Dickey-Fuller and Johansen) find in favour of random walks, whereas Phillips-Perron tests reject. Galbraith and Green (1992) looked at monthly data extending from 1880 to 1988 and do not find evidence of a random walk. They find that a variation of (2) plus a trend offered the best fit to the available data, yielding a long-run trend of 0.06 ^{o}C per decade.

A third approach pioneered by Fomby and Vogelsang (2000) uses a new trend estimator, robust to whether the model residuals have unit roots or not. They examine six series of annual global surface temperature anomalies. On data ending at or before 1997 they find a trend of 0.033 to 0.053 ^{o}C per decade, with the lowest value estimated on the data set with the longest time span (1851-1997). The highest trend is in the "Wilson-Hansen" data which end at 1993. The "JOBP" data base extends from 1856 to 1997 and yields a trend of 0.037 ^{o}C per decade. The CRU data examined below is very similar to JOBP, and is updated to the end of 2000.

The temperature series used in the next section do not seem to be random walks. The surface data from the CRU (Hadley Centre 2001) is the longest currently available global record with both land and sea coverage. The augmented Dickey-Fuller, Phillips-Perron, Bierans DHOAC, and Bierans-Guo tests[5] reject the null hypothesis of a unit root against a trend-model alternative, whereas the KPSS test finds the opposite. The satellite data, retrieved from the National Oceanic and Atmospheric Administration (NOAA 2001), span 1979:1 to 2001:1. The unit root hypothesis is rejected by all tests. Tests for unit roots in the balloon temperature data were rejected by all but the Augmented Dickey-Fuller statistic.

5. Temperature Trends in Global Datasets

The calculations reported here are based an ARMA technique, with details in the Appendix. Table 1 shows that the long-run trend in the surface data is 0.0445 ^{o}C/decade, with a 95% confidence interval no wider than 0.01 to 0.08 ^{o}C per decade and likely around 0.03 to 0.06 ^{o}C per decade. The standard error of the trend is approximately 0.0075. This means that if temperatures are to go up by the minimum of the IPCC projection envelope (1.4 ^{o}C), the trend per decade over the next century must rise by about 13 standard deviations above its level over the past century. If temperatures are to begin increasing by the central value of the current IPCC projection envelope (3.6 ^{o}C per century) then the decadal trend must rise over 42 standard deviations above its rate for the past century. These calculations do not rule out such future developments, they only tell us that if this change does occur it will be a remarkable deflection away from past behaviour.

We get closer to a testable hypothesis by looking at the data over the last two decades. It is clear in Figure 2 that after 1979 a distinct upward trend appears to emerge. But in focusing on the post-1979 period we are confronted with three problems. First, trends over a short span like two decades can be misleading. For instance, from 1910 to 1945 the linear trend is about 0.15 ^{o}C per decade, but that warming phase subsequently reversed into a cooling trend for the next 35 years. Hence the most recent 21 years is not long

---

[5] As implemented in Bierans (2000).

enough to establish a long-run climatology. Second, the very strong but temporary 1998 El Nino is an outlier which deflects the trend upward, but is apparently of natural origin and has since reversed out of the tropospheric data. Third, over this period the global satellite and surface data diverge.

Table 1 shows that in the post-1979 subsample the surface trend is approximately 0.15 ^{o}C per decade, with a 95% confidence interval of approximately (0.06, 0.24) ^{o}C per decade, placing it within the range of IPCC projections. By contrast, in the satellite MSU data there is an insignificant warming trend of only 0.020 ^{o}C per decade, in line with (indeed lower than) the long-run surface trend. The 95% confidence interval around the satellite data is approximately (-0.15, 0.19) ^{o}C per decade. In the tropospheric balloon data there is an insignificant trend of -0.015 ^{o}C per decade. That the balloons show no warming taking place, and that the trend is insignificant, corroborates the finding from the satellite data that the troposphere hasn't warmed in the past two decades. Unless there is a global meteorological mechanism that radiates more heat to the surface than is generated in the atmosphere, something other than tropospheric warming via an enhanced greenhouse effect is causing recent increases in measured surface temperatures. Urban heat island effects and the possibility that sea surface data falsely signal warming in marine air temperatures since 1979 are possible explanations, but the debate will no doubt take many years to settle.

The final source of data on global temperatures cannot be analyzed here using time-series methods, but does provide some insight into trends in specific locations. We are particularly interested in whether high-resolution proxies show a net warming pattern over the postwar period. Crowley and Lowery (2000) present 10 records that extend up to the present, from sites including the Western US, Alberta, the Sargasso Sea, Sweden, France, Germany, Siberia, China and Tibet. The German (Black Forest) tree ring record shows a marked increasing trend in the postwar period, as does the Chinese record, though the latter remains well below its early-twentieth century level. All the rest of the records show flat or declining temperature proxies after 1950. Jones et. al (1998) examine 17 proxy records from around the world. There is a net postwar proxy temperature increase only in records from the central UK, Tasmania and Antarctica. The rest show flat or declining trends. Briffa (2000) compiles nine tree ring records from northern high latitudes, of which only one, from Russia, shows a postwar increase, though another record from a nearby site shows a net decrease. While such surveys only show limited results from specific sites, they have the advantage that the records are usually from remote areas far from the influence of cities. A more complete survey, in Singer (2001), concludes that contemporary proxy records on balance do not confirm the warming trend found in the surface thermometer record.

6. Some Policy Implications

Deciding that the world is warming, and even that humans are responsible, does not on its own necessitate any particular policy response. There is still the question of costs and benefits of any proposed strategy. Nevertheless, the starting point for the discussion must be the basis of recent assertions that the world is experiencing an unusual warming. This paper has reviewed findings that are well known in the specialist literature but have not filtered into general policy discussions.

There is a long run trend in global surface data of approximately 0.4 ^{o}C per century, though analysts are divided over whether this is a statistical trend or a spurious feature in data that are actually following a random walk. Assuming the former for the time being, it is less than one-third the minimum level forecast by the IPCC (2001) for the current century. If the long-run rate of warming is to enter the range hypothesized to be now underway the trend coefficient must increase by about 13 standard deviations above its current level. If it is to attain the middle of the projected warming rates it must rise about 42 standard deviations. Surface data of the past two decades suggest such an acceleration has begun, but there are enough concerns about the quality of these data to make independent corroboration a necessity. Over the same period global temperatures as measured by weather satellites follow a trend slightly below the long run rate, and thus show no indications of acceleration. The network of balloon-borne radiosondes also show no signs of warming in the atmosphere just above the surface. In several recent compilations of temperature proxy records there is, likewise, no net global warming over in the postwar period.

Meanwhile there is no question that the concentration of carbon dioxide in the atmosphere is rising. Almost 40 percent of total historical CO_2 emissions have occurred since 1979 and the concentration of CO_2 in the atmosphere has risen 9 percent during that time. In the light of the absence of warming in data from balloons and satellites, it is worth remembering that the climate is very complicated and we still have a lot to learn about its behaviour. To the extent that we can infer something about the temperature-CO_2 relationship from actual recent and longer term trends, the data do not currently validate model projections of rapid warming as a result of rising greenhouse gas concentrations.

There are obvious policy implications to this assertion that hardly need to be stated. The Kyoto Protocol is a very expensive economic policy commitment. The decision to reduce and cap CO_2 emissions was made in the belief that they are causing global warming, and that global warming is a bad thing. If it turns out they are not causing warming, the policy is obviously not warranted. Even if they are causing warming, but only to a small degree, emission reductions are still likely unwarranted. As such, the fact that the evidence for global warming remains so ambiguous is extremely pertinent to the policy debate. Before any permanent policy commitments are undertaken it is imperative that these unresolved issues be subject to a much more thorough discussion.

References

Angell, J.K. (2000). “Global, Hemispheric, and Zonal Temperature Deviations Derived from Radiosonde Records.” In *Trends Online: A Compendium of Data on Global Change.* Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee, U.S.A.

Annan, Koffi (2001). “Sustainable Development: Humanity’s biggest challenge in the new century”. Speech delivered in Dhaka, Bangladesh, 14 March 2001. UN Press Release SG/SM/7739.

Balling, R.C. and W. S. Hughes (1996). “Urban Influences on South African Temperature Trends” *International Journal of Climatology* 16:935—940.

Bierens, H. J. (2000), "EasyReg 2000", Department of Economics, Pennsylvania State University, University Park, PA

Böhm, Reinhard (1998). “Urban Bias in Temperature Time Series—A Case Study for the City of Vienna, Austria.” *Climatic Change* 38:113-128.

Briffa, Keith R. (2000). “Annual Climate Variability in the Holocene: Interpreting the Message of Ancient Trees.” *Quaternary Science Reviews* 19: 87—105.

Campbell, Ian, Celina Campbell, Michael Apps, Nathaniel Rutter and Andrew Bush (1998). “Late Holocene ~1500 yr Climatic Periodicities and the Implications.” *Geology* 26 pp. 471—473.

Christy, John R., Roy W. Spencer and William D. Braswell (1997). “How Accurate are Satellite ‘Thermometers’?” *Nature* 389(25), p. 342

Christy, John R., David E. Parker, Simon J. Brown, Ian Macadam, Martin Stendel and William B. Norris (2001). “Differential Trends in Tropical Sea Surface and Atmospheric Temperatures since 1979.” *Geophysical Research Letters* 28(1), 183—186.

Crowley, Thomas and Thomas Lowery (2000). “How Warm was the Medieval Warm Period?” *Ambio* 29: 51—54.

Fomby, Thomas and Timothy Vogelsang (2000). “The Application of Size-Robust Trend Statistics to Global Warming Temperature Series.” Working Paper, Cornell University, September 2000.

Galbraith, John W. and Christopher Green (1992). “Inference about Trends in Global Temperature Data.” *Climatic Change* 22, 209-221.

Hadley Centre (1999). “Climate Change and its impacts: Stabilisation of CO_2 in the atmosphere.” U.K. Met.Office (http://www.met-office.gov.uk/sec5/CR_div/CoP5/).

Hadley Centre (2001). “Global Monthly Temperatures.” http://www.met-office.gov.uk/research/hadleycentre/CR_data/Monthly/HadCRUG.txt.

Hengeveld, Henry (2000). “Projections for Canada’s Climate Future.” *Climate Change Digest* CCD-00-01, Special Edition.

Hoyt, D.V. and K. H. Schatten, (1993). “A discussion of plausible solar irradiance variations, 1700-1992.” *Journal of Geophysical Research*, 98A, 18895-18906. Updated data supplied by D.V. Hoyt to the author by personal communication.

Hurrell, James W., Simon Brown, Kevin Trenberth and John R. Christy (2000). “Comparison of Tropospheric Temperatures from Radiosondes and Satellites: 1979-98.” *Bulletin of the American Meteorological Society* 81(9) 2165-2177.

Intergovernmental Panel on Climate Change (IPCC) Working Group I (1990) *Climate Change: The IPCC Scientific Assessment*. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC) Working Group I (1995) *Climate Change 1995: The IPCC Second Assessment Report*. Cambridge: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC) Working Group I (2001) *Climate Change 2000: The IPCC Third Assessment Report*. Cambridge: Cambridge University Press.

Jones, P.D., K.R. Briffa, T.P. Barnett and S.F.B. Tett (1998). “High-resolution Paleoclimatic Records for the Last Millennium: Interpretation, Integration and Comparison with General Circulation Model Control-run Temperatures.” *The Holocene* 8, 455—471.

Kärner, Olavi (2001). “On Stationarity and Antipersistency in Global Temperature Series.” Working Paper, Tartu Observatory, Estonia.

Kärner, Olavi (1996). “Global Temperature Deviations as a Random Walk.” *Journal of Climate* 9(13): 656—658.

Keeling, C.D. and T.P. Whorf 2000a. “The 1,800-year oceanic tidal cycle: A possible cause of rapid climate change.” *Proceedings of the National Academy of Sciences USA*, 97, 3814-19.

Keeling, C.D. and T.P. Whorf. 2000b. “Atmospheric CO_2 records from sites in the SIO air sampling network.” 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.

Magee, N., J. Curtis and N. Wendler (1999). “The Urban Heat Island Effect at Fairbanks, Alaska.” *Theoretical and Applied Climatology* 64, 39—47.

Marland, G., T.A. Boden, R. J. Andres, A. L. Brenkert, and C. Johnston. (1999). “Global, Regional, and National CO_2 Emissions.” 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.

Monasterski, Richard (1989). “Looking for Mr. Greenhouse.” *Science News* 135 pp. 216—221.

NRC (National Research Council), Panel on Reconciling Temperature Observations (2000) *Reconciling Observations of Global Temperature Change*. Washington: National Academy Press.

Peterson, T.C. and R.S. Vose (1997). “An Overview of the Global Historical Climatology Network Temperature Database” *Bulletin of the American Meteorological Society* 78:2837-2849.

Revkin, Andrew. “Endless Summer: Living with the Greenhouse Effect.” *Discover* October 1988, pp. 50—61.

Singer, S. Fred (1999a) *Hot Talk: Cold Science, Global Warming’s Unfinished Debate* 2nd ed. Oakland: The Independent Institute.

Singer, S. Fred (2001). “Is the Climate Warming?” Presentation at the First International Conference on Global Warming and the Next Ice Age, Dalhousie University, August 2001.

Spencer, Roy and Christy, John R., (1990). “Precise Monitoring of Global Temperature Trends from Satellites.” *Science* 247 pp. 1558-1562. Data updated at http://vortex.nsstc.uah.edu/data/msu/t2lt/t2ltglhmam.d

Stern, David and Robert Kaufman (1997). “Time Series Properties of Global Climate Variables: Detection and Attribution of Climate Change.” Australian National University Center for Resources and Environmental Studies Working Paper 9702.

Thompson, David W.J., John M. Wallace and Gabriele C. Hegerl (2000). “Annular Modes in the Extratropical Circulation. Part II: Trends.” *Journal of Climate* 13:1018-1036.

Vogelsang T.J. (1998). “Trend Function Hypothesis Testing in the Presence of Serial Correlation.” *Econometrica* 66:123—148.

Watson, Robert T. (2001). “Do Scientists Have Compelling Evidence of Global Warming?” *Insight on the News* March 12, 2001. http://www.insightmag.com/archive/200103143.shtml.

Woodward, W. A. and Gray, H. L. (1993), “Global Warming and the Problem of Testing for Trend in Time Series Data,” *Journal of Climate*, 6, 953-962.

Zheng, Xiaogu and Reid E. Basher (1999). "Structural Time Series Models and Trend Detection in Global and Regional Temperature Series." *Journal of Climate* 12, 2347-2358.

TABLE 1
Trend Estimation Results for Three Different Data Sources
MONTHLY, 1856.1 to 2000.12

Data Source	Time Span	Trend/ Decade	Standard Error	95% Confidence Interval	Estimating Equation
CRU Surface Series (monthly)	1856-2001	0.045 °C	0.0075	0.03—0.06	ARMAX(5,3)
CRU Surface Series (monthly)	1979-2001	0.150 °C	0.045	0.06—0.24	ARMAX(2.0)
Satellite MSU (monthly)	1979-2001	0.020 °C	0.085	-0.15—0.19	ARMAX(2.0)
Balloon Radiosonde (monthly)	1979-2001	-0.015 °C			ARMAX(1,0)

For definitions of data sources see text. For computational details see Appendix.

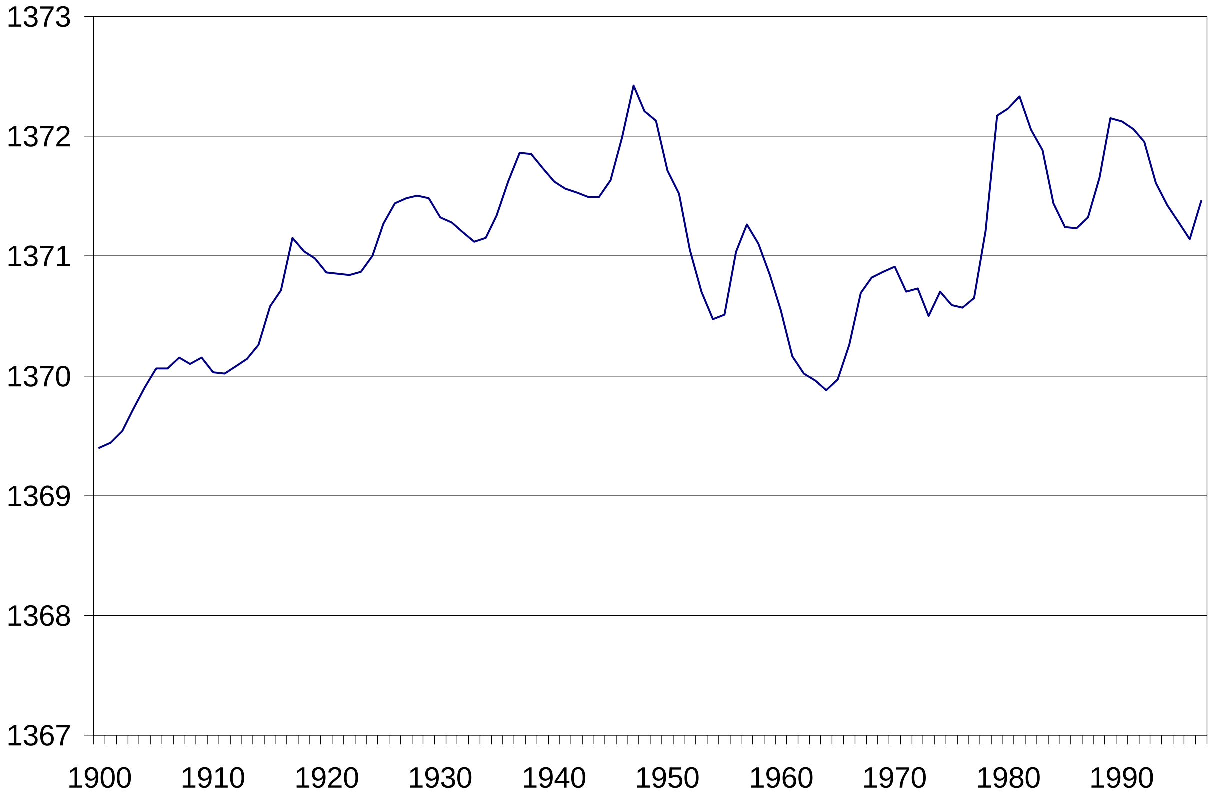

Figure 1: Annual Solar Irradiance since 1900 in watts per square meter (Hoyt and Schatten 1993).

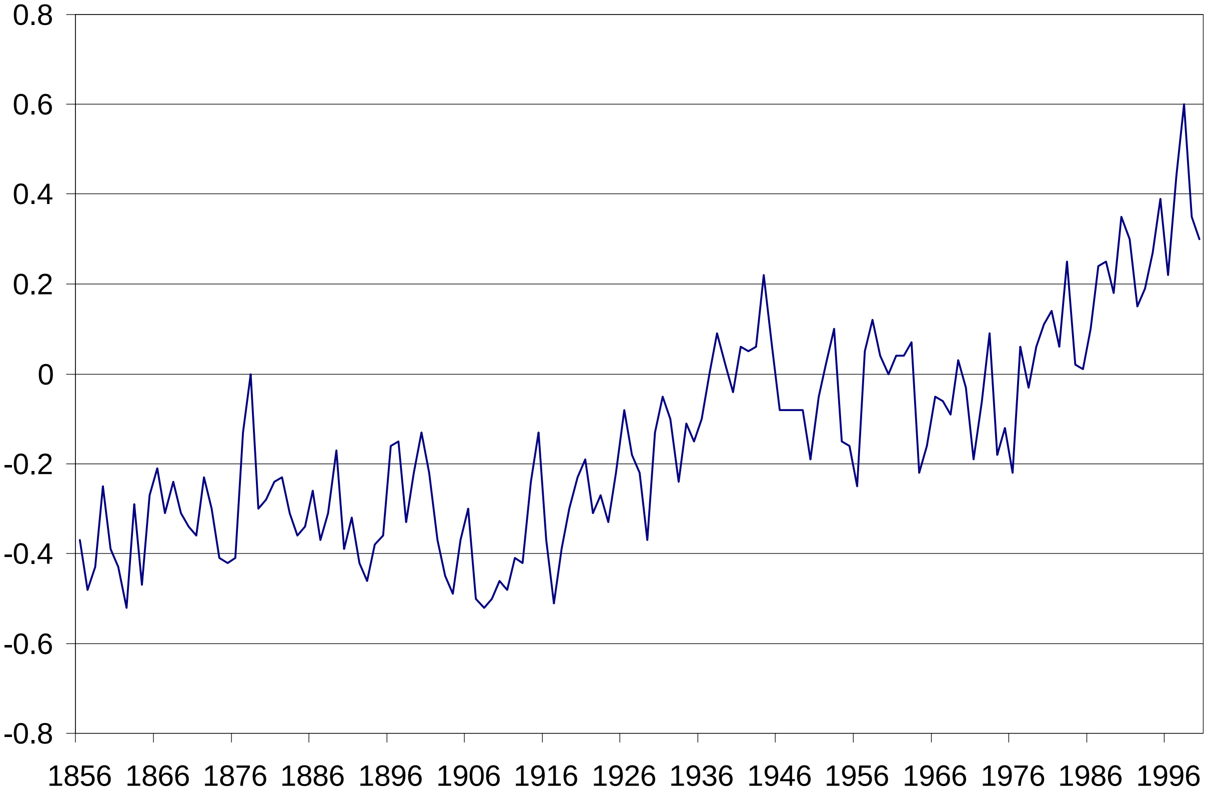

Figure 2: Annual global surface temperature anomalies (^{o}C) from 1856 to 2000, relative to a 1961-1990 mean. Source: Hadley Centre (2001).

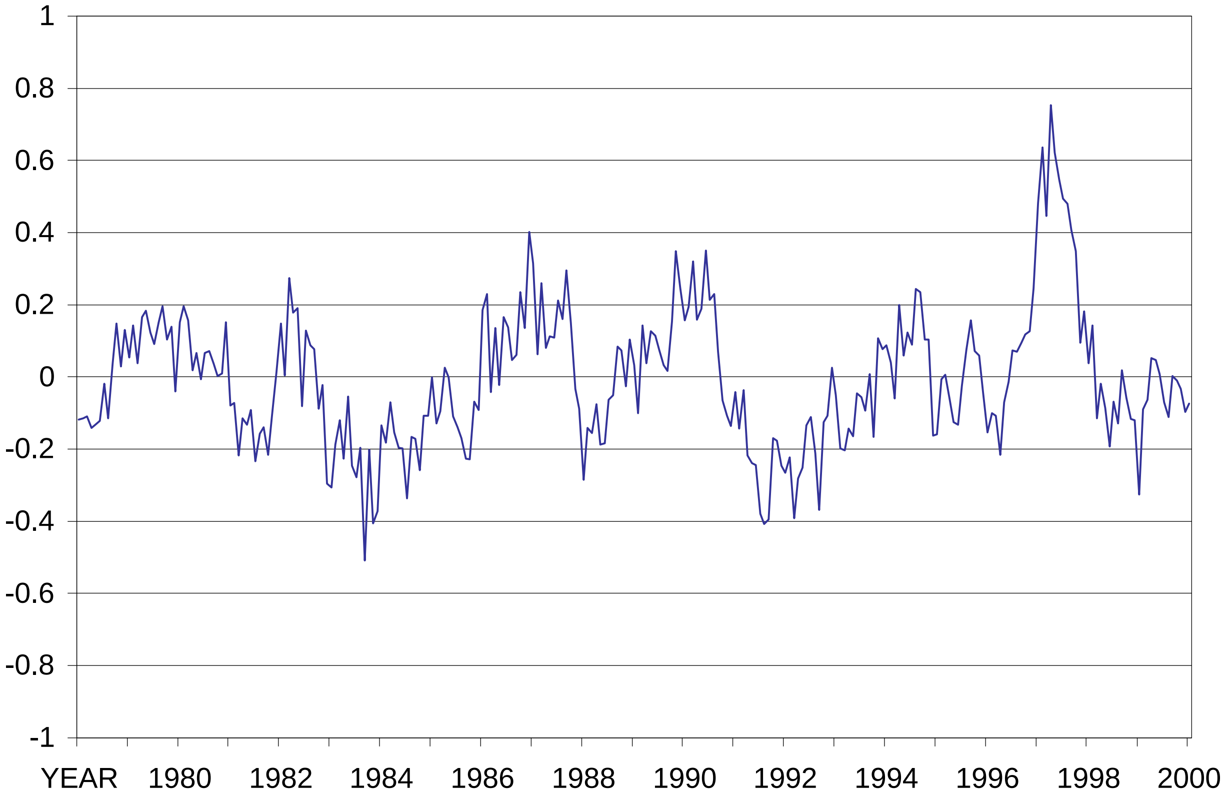

Figure 3: Monthly global temperature anomalies (°C), January 1979-January 2001 as measured by NOAA weather satellites, relative to 1979-2000 mean. Source: Spencer and Christy (1990)

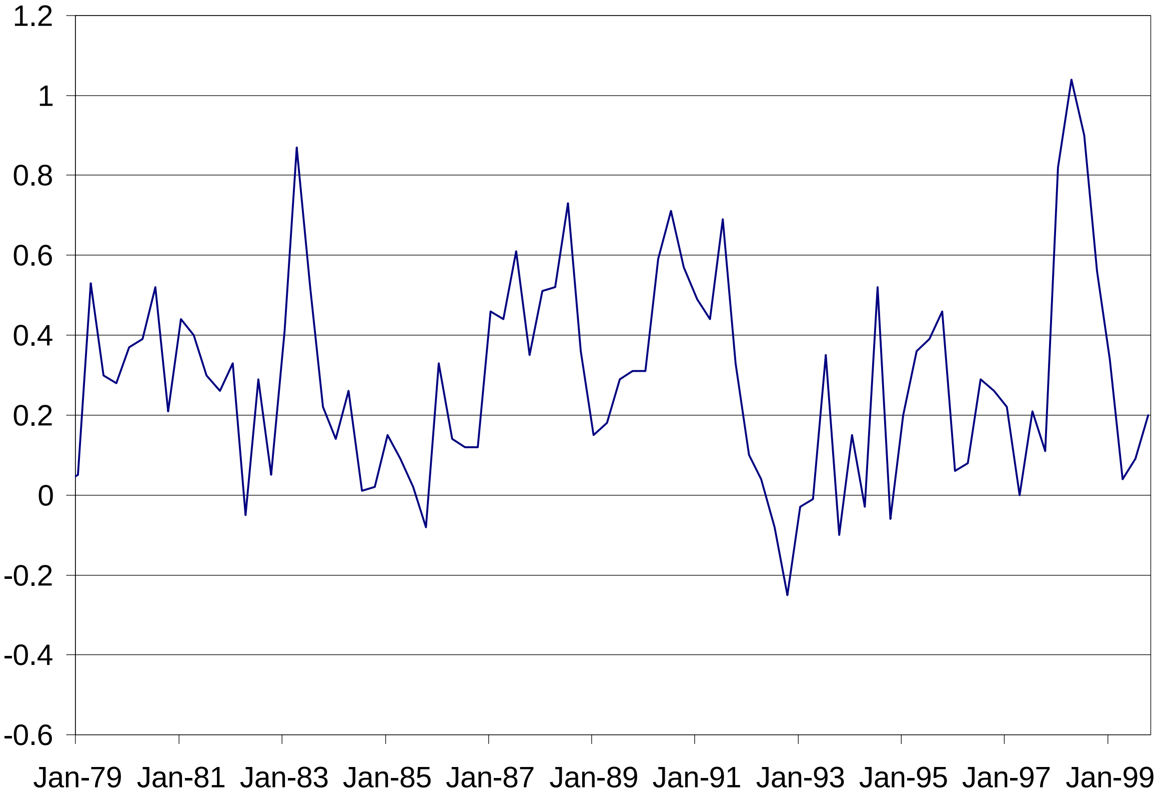

Figure 4: Quarterly global temperature anomalies (^{o}C) 1979.1-1999.4 as measured by balloon-borne radiosondes (850—300 mb layer), relative to 1958-1977 mean. Source: Angell (2000)

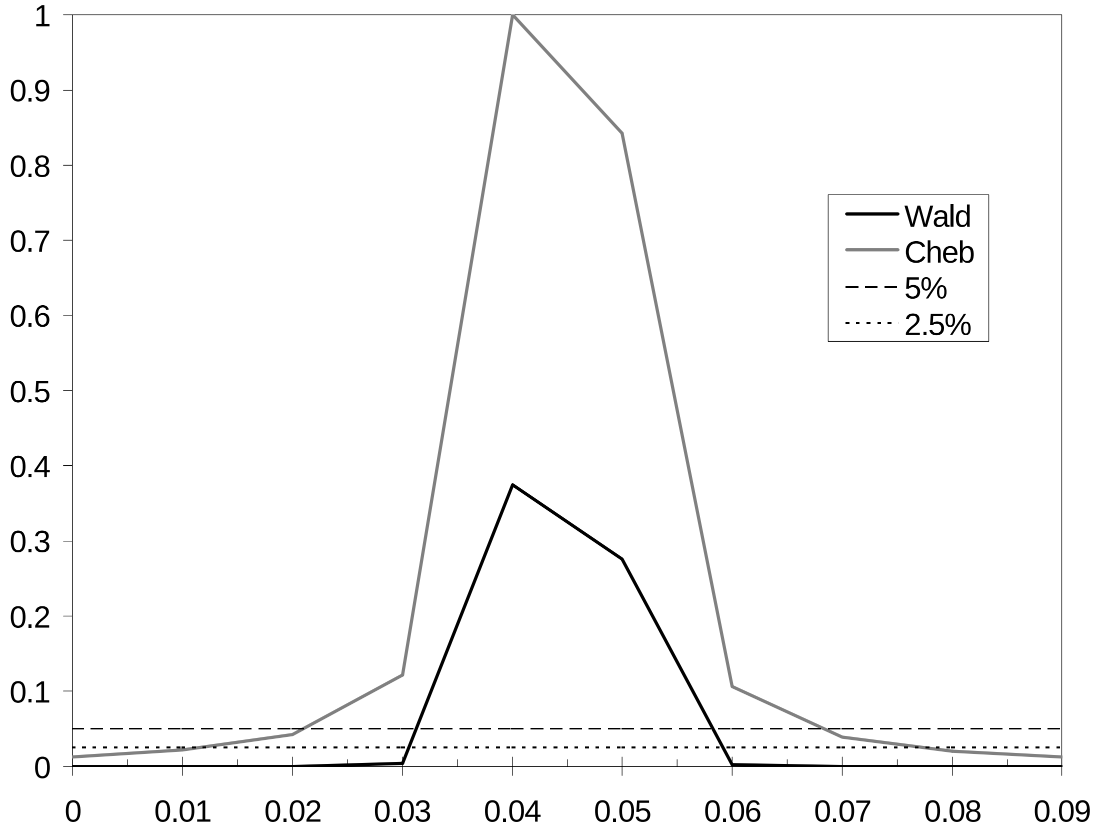

Figure 5: P-values for hypothesis tests of trend values in monthly surface data, 1856-2000. “Wald” denotes P-value for Wald test which follows asymptotic c^2 distribution. “Cheb” denotes P-value computed from Chebyshev’s inequality, and represents maximum probability of accepting the null hypothesis for any distribution of the trend estimator. 5% and 2.5% denote significant levels.

Appendix: Statistical Estimation

Table 1 shows the trend estimator from monthly surface anomaly data which extend from January 1856 to December 2000. The dimensions for this and all subsequent ARMA models were selected by fitting an ARMA(6,6) equation and testing for joint significance of higher-order terms in groups until further reductions were rejected. The surface data model is ARMA(5,3):

$$a_t = q_0 + q_1 T + \sum_{i=1}^{5} f_i a_{t-i} + e_t + \sum_{j=1}^{3} d_j e_{t-j} \ .$$

All estimations were done using the econometrics software SHAZAM. The instantaneous trend parameter q_1 is transformed into the long-run trend using $q_1 / (1 - \sum f_i)$, yielding 0.0445 ^{o}C per decade, or about 0.4 degrees per century, in good accord with other estimates using other estimating techniques.

A confidence interval around the long-run trend was calculated using the nonlinear hypothesis test option in SHAZAM. A Wald statistic with an asymptotic c^2 distribution yields a P-value which drops below 0.025 for a test of [trend $\leq$ 0.031/decade] and for a test of [trend $\geq$ 0.059] per decade, so the trend can be bounded within the approximate interval (0.03, 0.06) ^{o}C per decade. Since the small-sample distribution of this test statistic is not known, Chebyshev's inequality (Rice 1988) can be used to provide maximum bounds on the confidence interval. These are valid in small samples but represent the outer limits, so the true confidence interval is likely narrower. The P-value associated with Chebyshev's inequality drops below 0.025 for [trend $\leq$ 0.012/decade] and for [trend $\geq$ 0.077/decade], so the 95% confidence interval on the long-run trend is no wider than (0.01, 0.08) ^{o}C per decade. Figure 5 is a graph of the P-values associated with the null hypothesis of each trend value on the horizontal axis.

We can approximate the standard error of the trend using one-quarter of the 95% confidence interval defined by the Wald chi-squared statistics, which equals 0.0075. If temperatures are to begin going up by 1.4 ^{o}C/decade the long term decadal trend must rise by about (0.14-0.0445)/0.0075 = 12.7 standard deviations from its current level.

The post-1979 surface data subsample trend is computed using an ARMA(2,0) regression with trend. The confidence interval was calculated as above, using chi-squared and Chebyshev statistics. The Wald statistic yields a 95% confidence interval of (0.06, 0.24) ^{o}C per decade while the Chebyshev outer bounds are much wider, approximately –0.06 to +0.40. The satellite MSU trend was estimated using ARMA(2,0). There is an insignificant warming trend of 0.020 ^{o}C per decade, the 95% confidence interval using the Wald statistic is approximately (-0.15, 0.19) ^{o}C per decade and the Chebyshev outer bounds are considerably wider (at least ± 0.25). An ARMA(1,0) model was applied to the balloon-borne radiosonde data over the period 1979:1 to 1999:4. There is a (transformed) trend of -0.015 ^{o}C per decade. The trend is insignificant by the chi-squared statistic (P=0.8546), but the confidence intervals were not computed since the trend is negative anyway.