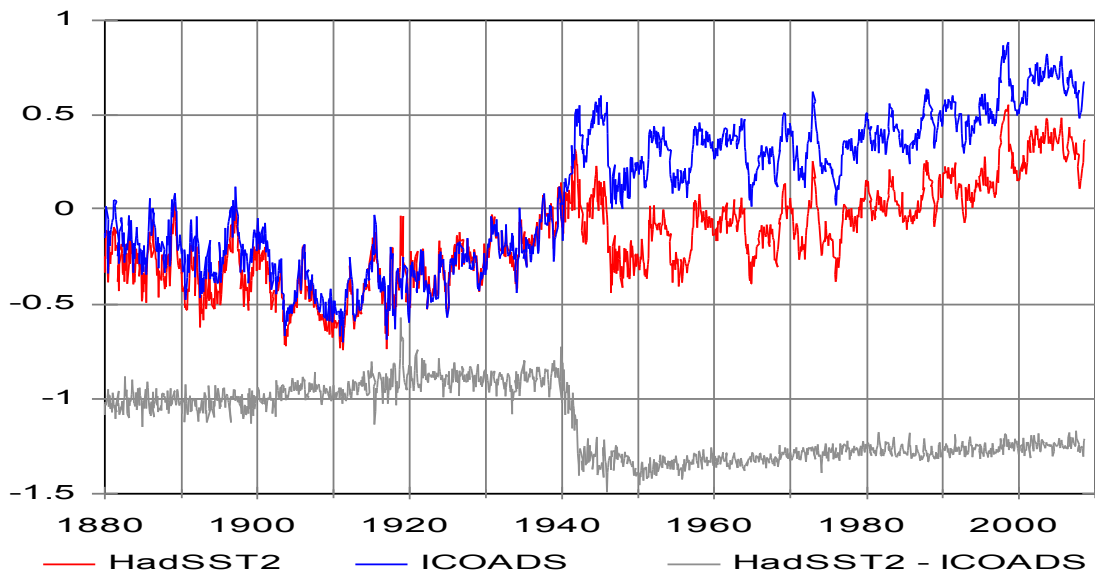


SST SERIES, AIR TEMPERATURE SERIES, AND THE WORLD WAR II SHIFT ADJUSTMENT

INTRODUCTION

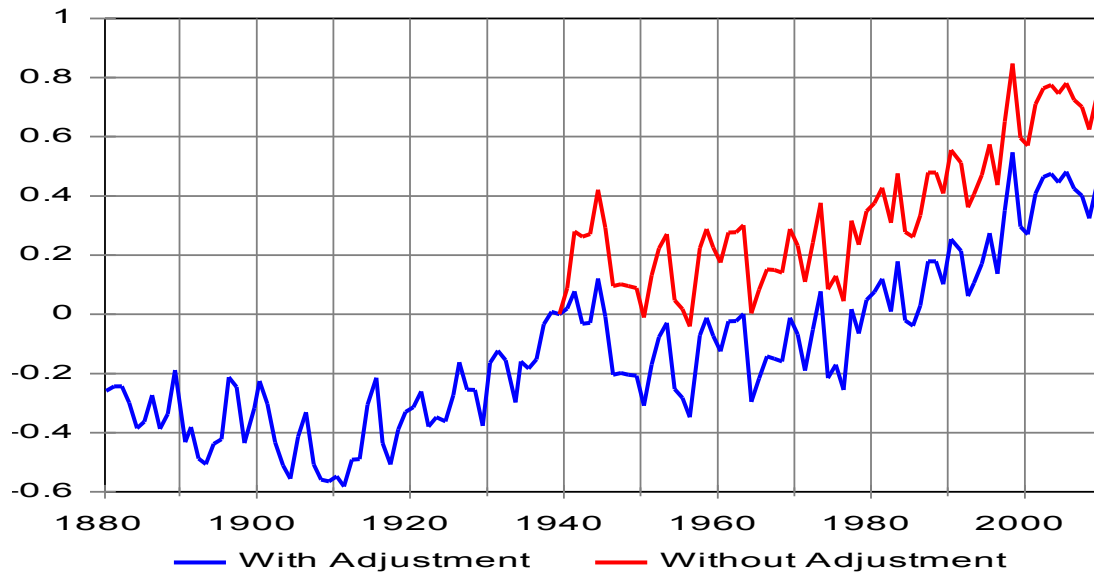
Various adjustments are applied to the SST record to remove biases thought to have caused by changes in the way SST measurements have been taken. But by far the largest single adjustment is the stair-step cooling adjustment that compensates for an artificial upward shift in the SST record thought to have been caused by a change from dominant bucket to dominant ship-intake SST measurements early in World War II. The size and timing of this adjustment is illustrated in Figure 1, which compares the current (July 2010) version of the unadjusted ICOADS SST series ⁽¹⁾ with the current version of the adjusted UK Met Office HadSST2 series ⁽²⁾. It shifts the entire ICOADS SST series after 1942 down by almost 0.5C relative to the series before 1940.

FIGURE 1
UNADJUSTED ICOADS VS. ADJUSTED HADSST2 SST SERIES (Degrees C, monthly means)



The World War II shift adjustment has a large impact on our perceptions of “global warming”. In fact, were it not applied the HadCRUT3 series the IPCC used to define surface air temperature trends in the 2007 AR4 would show about 0.3C more warming. This is illustrated in Figure 2, which shows what the current version of HadCRUT3 ⁽²⁾ – which is substantially the same as the AR4 version – looks like with and without the adjustment. (Note: HadCRUT3 is an area-weighted average of the HadSST2 “ocean” series and the CRUTEM3 “land” series, and because oceans cover most of the earth it is based 60-70% on HadSST2 and only 30-40% on CRUTEM3. It has therefore been assumed that removing the shift adjustment from HadSST2 would shift HadCRUT3 by $0.6 \times 0.5\text{C} = 0.3\text{C}$ upwards after 1942. This assumption is not exact but is acceptable for illustration purposes.)

FIGURE 2
HadCRUT3 WITH AND WITHOUT WWII SHIFT ADJUSTMENT (Degrees C, monthly means)



It is obviously important to make sure that an adjustment that has this large an impact is valid. For it to be valid the following conditions must be met:

1. Ship-intake SST measurements are biased at least 0.5C high overall relative to bucket SST measurements (or a bucket-intake change could not have caused a 0.5C shift in the SST record).
2. There was a permanent change from dominant bucket SST measurements to dominant ship-intake SST measurements early in World War II.
3. There was an artificial and permanent upward shift in the SST record that coincides with the bucket-intake change.

The question of whether these conditions are met addressed in sequence below.

(Notes: The Figures in this review are constructed using data downloaded from the internet, with web addresses provided as references. Studies that are not available on the web are referenced indirectly through other studies that describe their results, e.g. ^(in 11). Numeric data are used where available, but in cases where they are not values have been scaled off graphs. This will have led to some imprecisions, but these should not be large enough to affect the conclusions. All series are global unless otherwise specified. Finally, the term “adjustments” is used in preference to “corrections” because “corrections” implies that the adjusted data are more accurate than the unadjusted data, which is often not the case.)

Condition 1: Is the bucket-intake bias large enough to generate a 0.5C shift?

Estimates of the amplitude of the bucket-intake bias made over the last fifty years are summarized in Table 1:

TABLE 1
ESTIMATES OF AMPLITUDE OF BUCKET-INTAKE BIAS

Study	Area studied	Data from years	No. of Samples	Intake-bucket bias	Comments
Saur 1963 ⁽³⁾	N. Pacific	1960-62	6,826	0.7C (1.2F)	Custom-designed bucket, US navy ships
Roll 1965 ⁽⁴⁾	?	?	5,689	0.18C	Collocated readings
Walden 1966 ⁽⁵⁾	?	?	13,847	0.3C	German ships
James & Fox 1972 (in 11)	Global	1968-1970	16,000	0.3C	Collocated readings
Collins et al. 1975 (in 11)	N Pacific	1927-1933	?	0.3C	Canadian intakes vs. Japanese buckets
Tabata 1978 ⁽⁶⁾	N Pacific	1974-1976	1,150(?)	0.2C	Intakes(?) vs. buoys
Barnett 1984 ⁽⁷⁾	N Pacific	“long-term”	~10,000	0.4C	Intakes vs. bathythermographs
Kent et al. 1991 ⁽⁸⁾	N Atlantic	1988-1990	22,592	0.3C	Multinational ships
Kent & Kaplan 2006 ⁽⁹⁾	N Atlantic	1975-1979	8,410	0.09C	Multinational ships
		1980-1984	11,245	0.15C	
		1985-1989	11,073	0.18C	
		1990-1994	5,122	minus 0.13C	
Kennedy et al. 2008 ⁽¹⁰⁾	?	1970-2004	?	0.09C	Collocated readings

If a bucket-intake change indeed caused a permanent 0.5C upward shift in the SST record after WWII then intakes must be biased at least 0.5C high relative to buckets, but the 0.7C Saur estimate at the top of the Table is the only one that exceeds this threshold. Most of the later estimates are in the 0.2-0.4C range, enough to explain only about half of the shift. And according to the most recent estimates (Kent & Kaplan, Kennedy) a change from 100% buckets to 100% intakes could by itself have caused an upward shift of only about 0.1C in the SST record after 1970.

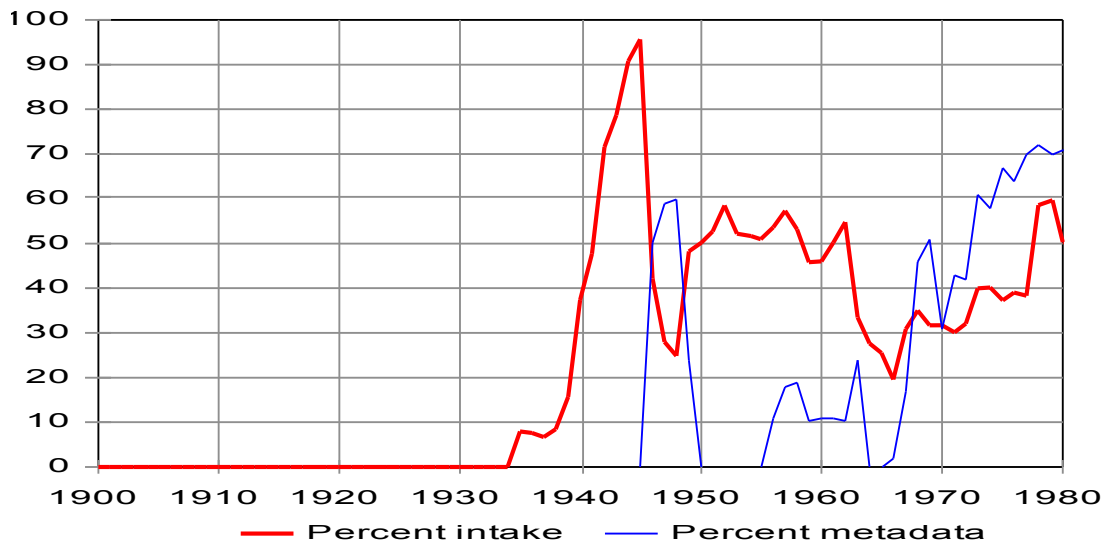
All the Table 2 estimates are subject to uncertainty because of small sample size (there are over 200 million individual SST readings in the ICOADS data base), short comparison periods, limited areal coverage and, in the case of the pre-1991 studies, a lack of metadata (these studies appear to have been based largely on assumptions as to which readings were taken with buckets and which with intakes). But despite these uncertainties the Table 2 estimates strongly suggest that the amplitude of the bucket-intake bias is not large enough for a bucket-intake change to have caused a permanent 0.5C shift in the SST record.

In other words, Condition 1 is not met.

Condition 2: Was there a change from dominant bucket to dominant intake SST measurements early in World War II?

Figure 3 shows the most recent (Kent et al. 2010, hereafter K10) ⁽¹¹⁾ estimates of intake measurement contributions as a percentage of total SST measurements between 1900 and 1980, when measurements were effectively either intake or bucket (data from K10 Figure 3). According to these results intake reading percentages increased from zero or near-zero before World War II to an average of about 40% after. However, the only time intake readings actually dominated was during World War II.

FIGURE 3
KENT ET AL. PERCENT INTAKE AND METADATA ESTIMATES (Percent, annual means)

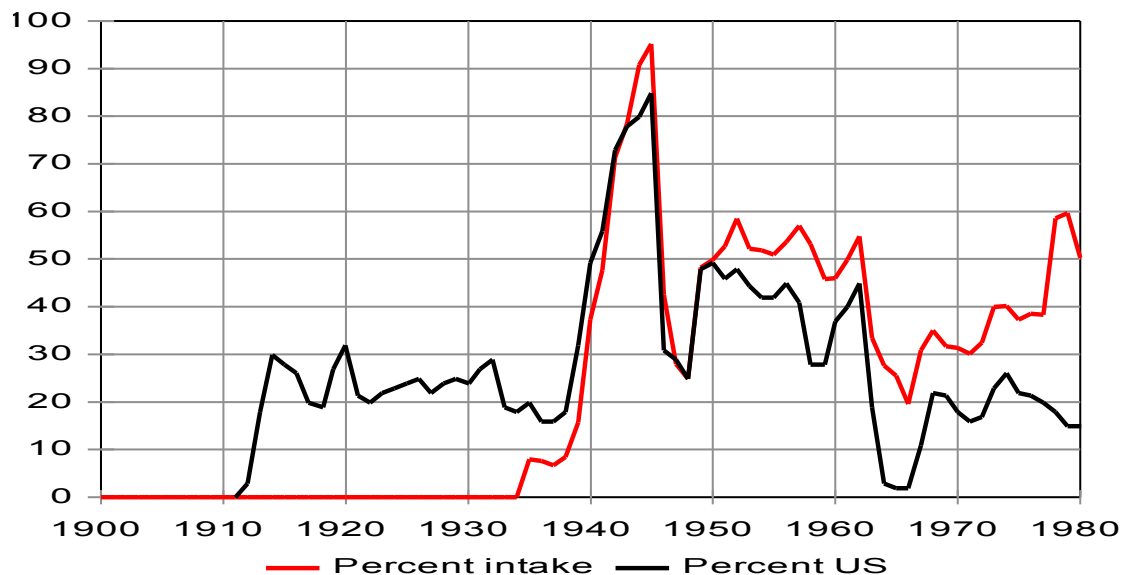


There are also reasons to believe that the pre-postwar bucket-intake change was not as large as Figure 3 suggests. To make robust estimates of the amplitude of the change we need metadata – i.e. records of how individual SST measurements were taken – but the percent metadata estimates (blue line, data again from K10 Figure 3) show that we have only sporadic metadata between 1946 and 1967 and none at all before 1946. Because of this the Figure 3 intake percentages are based largely on assumptions, with the main ones being 1) that US ships took dominant intake readings and non-US ships dominant bucket readings (with a gradual transition to intakes) after World War II, and 2) that essentially all readings before World War II were taken with buckets.

Assumption 1 is reasonable but Assumption 2 is not. Intake readings were undoubtedly taken before World War II. Folland and Parker (1995) ⁽¹²⁾ acknowledge that “in the 1920s and 1930s engine-inlet temperatures were taken on many ships”. Rayner et al. (2006) ⁽¹³⁾ state that “it appears that the United States started the switch to using engine room intake measurements as early as 1920”. It is known that Canadian ships took intake measurements in 1927 (Collins, Table 1) and that British ships took them in 1924 (Brooks 1926) ⁽¹⁴⁾. There are also reasons to believe that intake measurements may have been preferred. According to the 1932 Bulletin of the US National Resources Council ⁽¹⁵⁾; “The best modern practice on commercial ships is to expose the bulb of a sea water thermograph ... in a well projecting into the condenser intake pipe ... ” Finally, K10 acknowledge that “on ships with engines, it was convenient to measure the temperature of the pumped seawater used to cool the engines”.

In the absence of metadata there is no good way of estimating what prewar intake percentages were, but they were obviously greater than zero, meaning that Figure 1 will overstate the magnitude of the prewar-postwar bucket-intake change. In fact, if we use percent US measurements as intake percentage proxies both before and after World War II we find that the pre-postwar increase in intake percentages was probably only on the order of 10-20%. This is illustrated in Figure 4, which compares the Kent et al. intake percentages with US measurements as a percentage of total measurements (data from Thompson et al. 2008, Figure 4) ⁽¹⁶⁾.

FIGURE 4
KENT PERCENT INTAKE VS. THOMPSON PERCENT US (Percent, annual means)

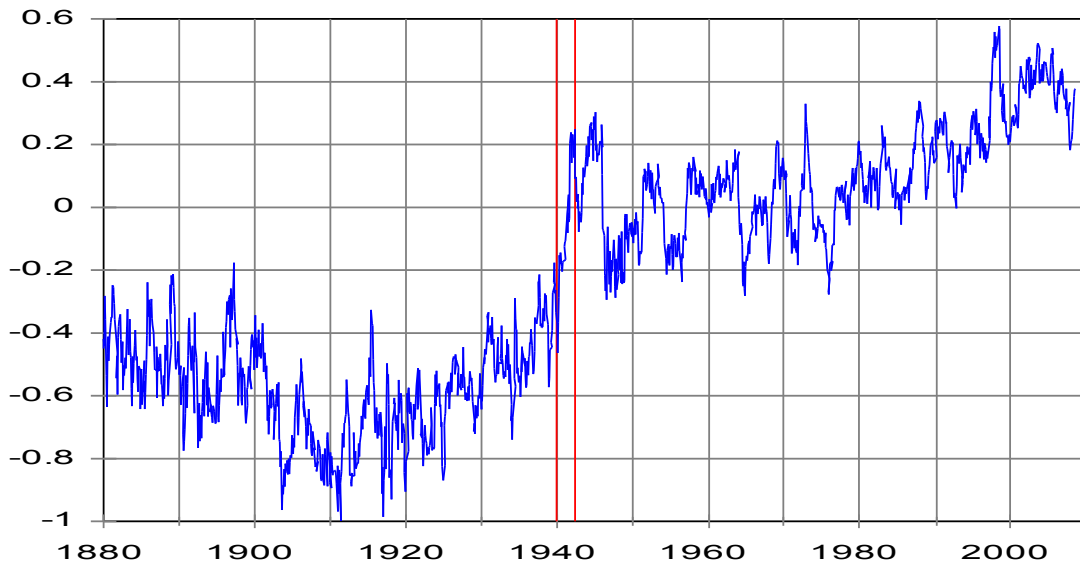


In summary, we do not have the metadata we need to determine whether there was or was not a bucket-intake change during World War II (as K10 put it: “Lack of metadata is an important contributor to the uncertainty in SST variations over time”) but the data we do have strongly suggest that there was no permanent change from dominant bucket measurements to dominant intake measurements after World War II. Therefore Condition 2 is not met either.

Condition 3. Was there an artificial shift in the SST record that coincides with the bucket-intake change?

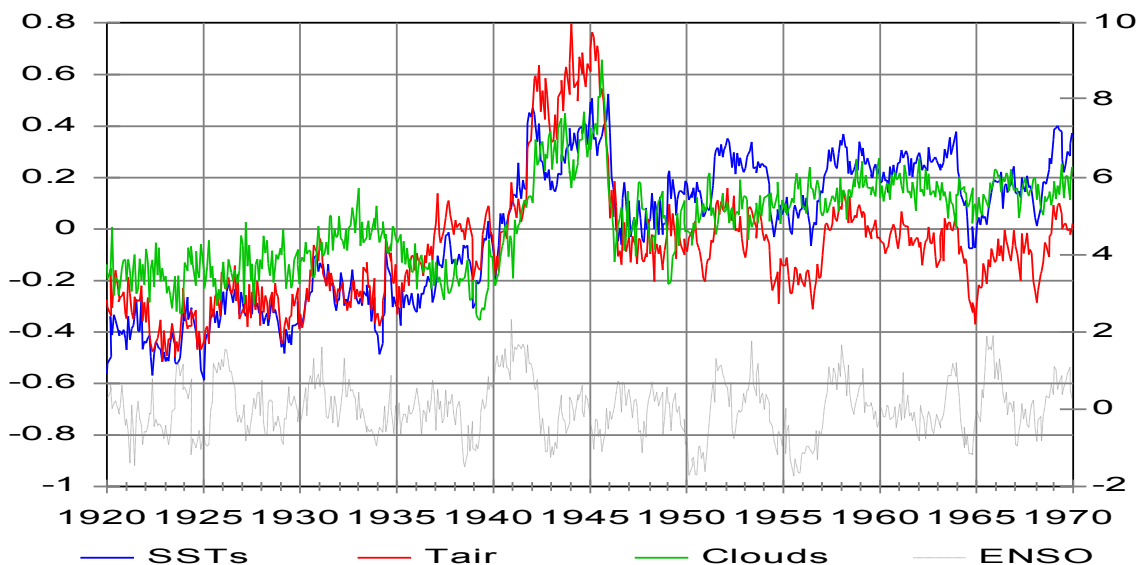
Figure 5 presents an expanded version of the ICOADS SST series shown earlier in Figure 1. There is no good evidence for a permanent upward shift coinciding with the bucket-intake change. There is a ragged upward shift that coincides with the period over which the bucket-intake change supposedly occurred (December 1939 to April 1942, defined by the vertical red lines) but it is promptly canceled out by an abrupt downward shift of roughly the same magnitude in 1945/1946. These two shifts form a “spike” that lasted only for the duration of World War II, plus a few months, and which causes no visible displacement in the SST record.

FIGURE 5
UNDAJUSTED ICOADS SST SERIES (Degrees C, monthly means)



Another interesting feature of the wartime SST spike is that it coincides with spikes in the ICOADS marine air temperature (Tair) and cloud cover series ⁽¹⁷⁾ (Figure 6). The chances that all three series would have spiked naturally during World War II are vanishingly small, so we can be confident that the three spikes are spurious. (The right-scale plot of the bivariate ENSO index ⁽¹⁸⁾ shows that they are not related to the 1940-41 El Niño.)

FIGURE 6
ICOADS SST, MAT AND CLOUD COVER SERIES (Degrees C, monthly means)

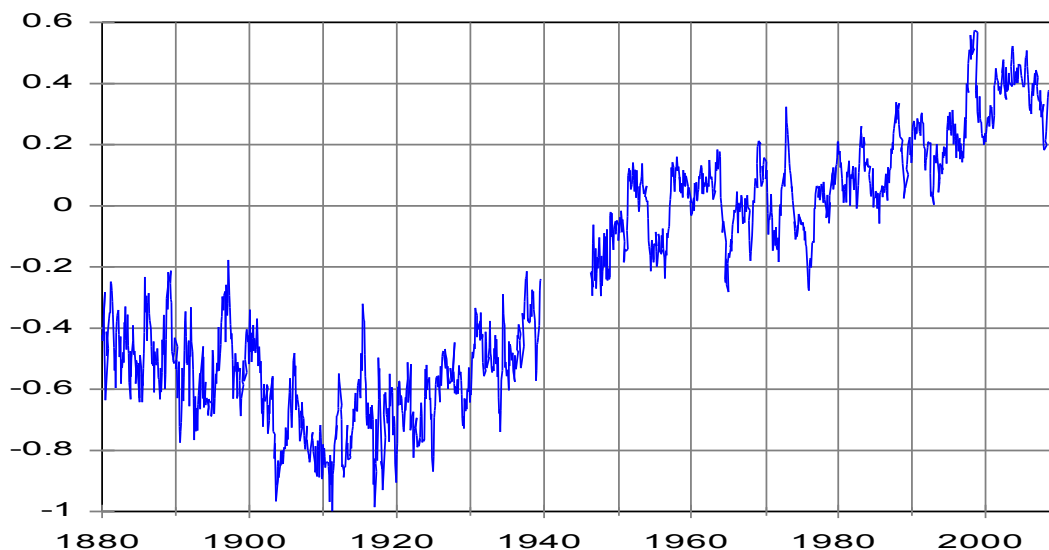


(The coincidence of these three spikes raises the question of whether the spike in the SST series was even caused by bucket-intake changes, since such changes obviously could not have caused the spikes in the MAT and cloud cover series. It has been speculated that taking night air temperature readings

inside the wheelhouse to avoid showing lights may have contributed to the MAT spike, but Tair shows that air temperatures did not return to “normal” levels until 1946, well after the need to darken ship had passed. The spike in the cloud cover series is unexplained. The most plausible explanation for the three spikes is in fact an across-the-board bias generated by wartime conditions, which among other things caused major changes in observational practices, a large decrease in the number of marine readings (SST readings fell from 824,600 in 1938 to 113,200 in 1942), a redistribution of marine readings from the Atlantic and Indian Oceans to the Pacific and from the extratropics to the tropics, and a wholesale change from non-US to US data sources. It is not known which of these changes had the largest impact, but the spikes correlate best with the wartime spike in the percentage of US readings (Figure 4), and the downward shifts in late 1945-early 1946 coincide with the demobilization of the US Navy.)

But regardless of what caused the spikes the SST data in and around World II are unquestionably bad, and this makes it acceptable to delete them. Deleting the data between September 1939 and April 1946 from the Figure 5 ICOADS series gives the results shown in Figure 7. There is no longer any sign of a permanent shift during World War II; all we see is forty or fifty years of continuous warming with no sign of a break in the trend. (Note that the “spikes” in the SST and Tair records around 1950 and in the late 1950s-early 1960s tend to coincide with El Niño events (Figure 6) and can therefore be assumed to be natural.)

FIGURE 7
ICOADS SST SERIES, NO WARTIME DATA (Degrees C, monthly means)



However, the main line of evidence for a permanent upward shift in the SST record is supposedly the large upward offset in the SSTs relative to the night marine air temperature (NMAT) series that occurs early in World War II. If there is such an offset it should be plainly visible in Figure 8, which compares the ICOADS SST series with the current version of the Rayner (2003) MOHMAT series ⁽¹⁹⁾ – the most commonly-used NMAT series – and plots the difference between the two. (Note that MOHMAT also shows the spurious wartime spike.) The difference plot does show an abrupt upward shift in the SSTs relative to the NMATs around 1940, but it is unclear whether it is a permanent offset or just a “blip” in a long-term differential warming trend. Another problem is that the offset begins in 1938, over a year before the bucket-intake change that allegedly caused it.

FIGURE 8
ICOADS SST SERIES VS. MOHMAT NMAT SERIES (Degrees C, annual means)

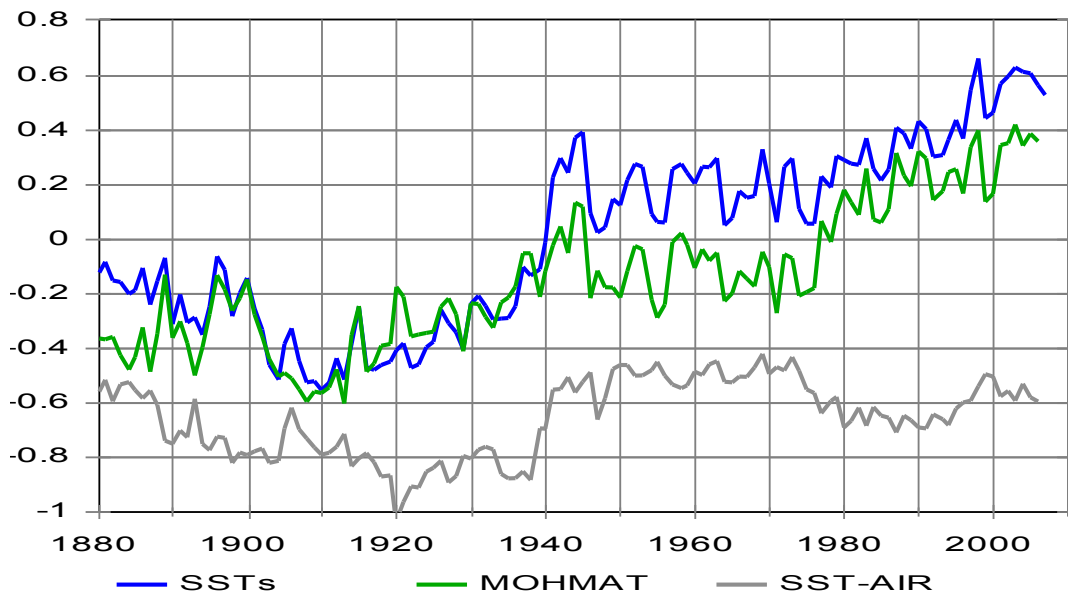


Figure 9 shows the SST minus NMAT plots for three other versions of the NMAT series – HadMAT and the Ishii et al. “uninterpolated” and “interpolated” series, with the wartime spikes deleted ⁽²²⁾. Like the Figure 6 difference plot they all show the SSTs warming by about 0.5C relative to the NMATs between 1920 and 1950-60, but none shows an abrupt wartime offset.

FIGURE 9
ICOADS MINUS OTHER NMAT SERIES, NO WARTIME DATA (Degrees C, annual means)

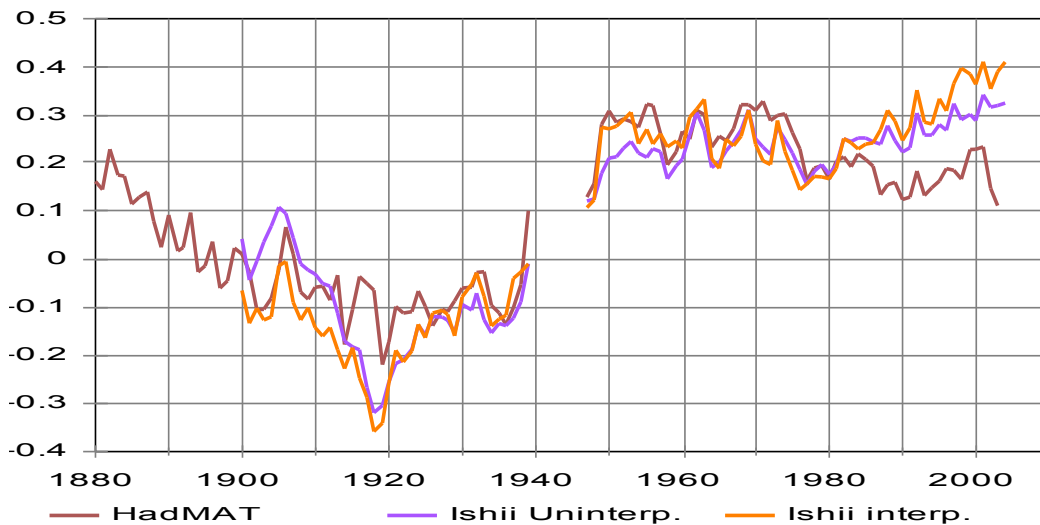
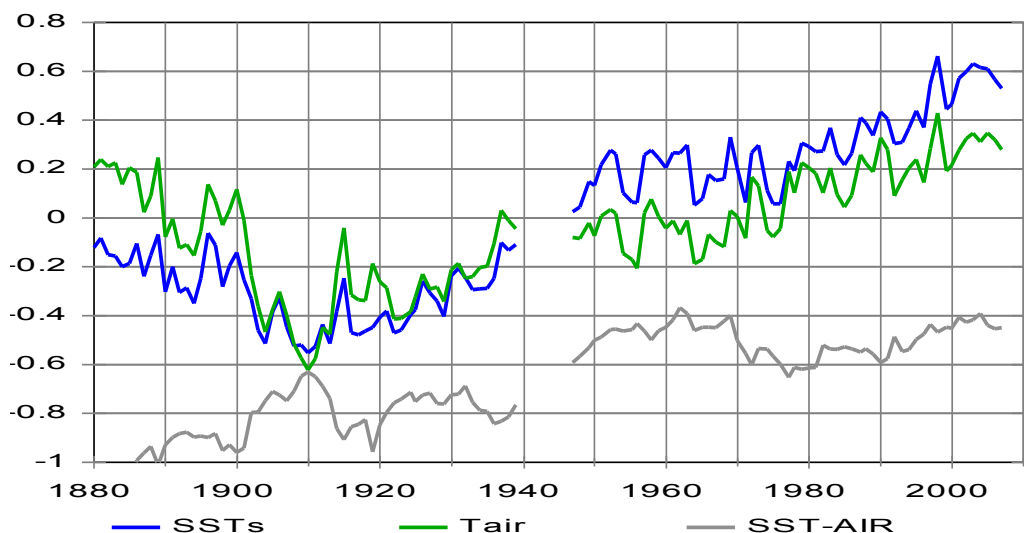


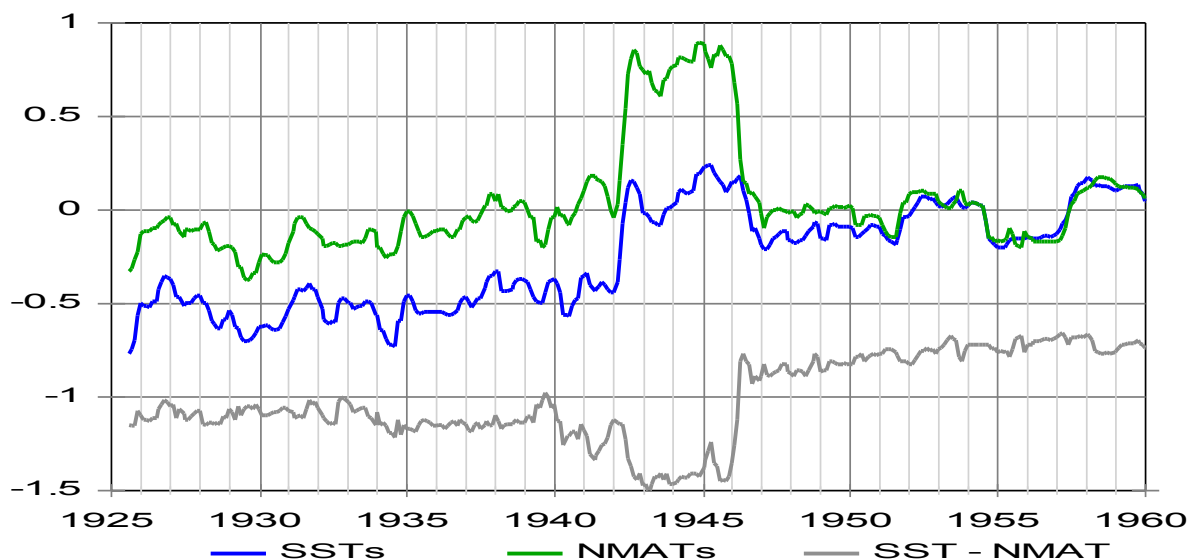
Figure 10 now compares the ICOADS SST series with the ICOADS Tair marine air temperature series shown in Figure 6. This is arguably a more robust comparison because Tair includes the daytime air temperature readings that the NMAT data sets exclude (they are thought to be contaminated by solar heating impacts) and because so far as is known Tair is unadjusted (NMAT series are adjusted for “deck height” biases). Again there is no sign of an abrupt wartime offset.

FIGURE 10
ICOADS SST VS. ICOADS Tair SERIES, NO WARTIME DATA (Degrees C, annual means)



Finally, Figure 11 shows the SST and NMAT series of Folland, Parker and Kates (1984, data from their Figure 2) ⁽²⁰⁾, who were the first to identify the World War II shift. They attributed it to a “sudden but undocumented change ... from the predominant use of canvas and other uninsulated buckets to the use of engine intakes” when the US entered World War II in December 1941, and applied a 0.3C stair-step cooling adjustment to remove it. However, their series also showed no evidence for this shift. They show large and obviously spurious wartime spikes in the SST and NMAT series but no sign of 0.3C of pre-postwar displacement in the SST series. The SST minus NMAT plot also shows no evidence for a permanent upward offset, and in fact it shows a downward offset between 1941 and 1942 when the upward shift in the SST record is supposed to have occurred. The upward offset does not occur until 1945-46.

FIGURE 11
FOLLAND, PARKER & KATES 1984 SST AND NMAT SERIES (Degrees C, monthly means)



These results demonstrate that there was no abrupt and permanent upward shift in the SST series nor any abrupt and permanent upward offset between the SST series and the NMAT or MAT series early in World War II. This, of course, does not prove that the SST record is unbiased over this period, but Condition 3 is clearly not met.

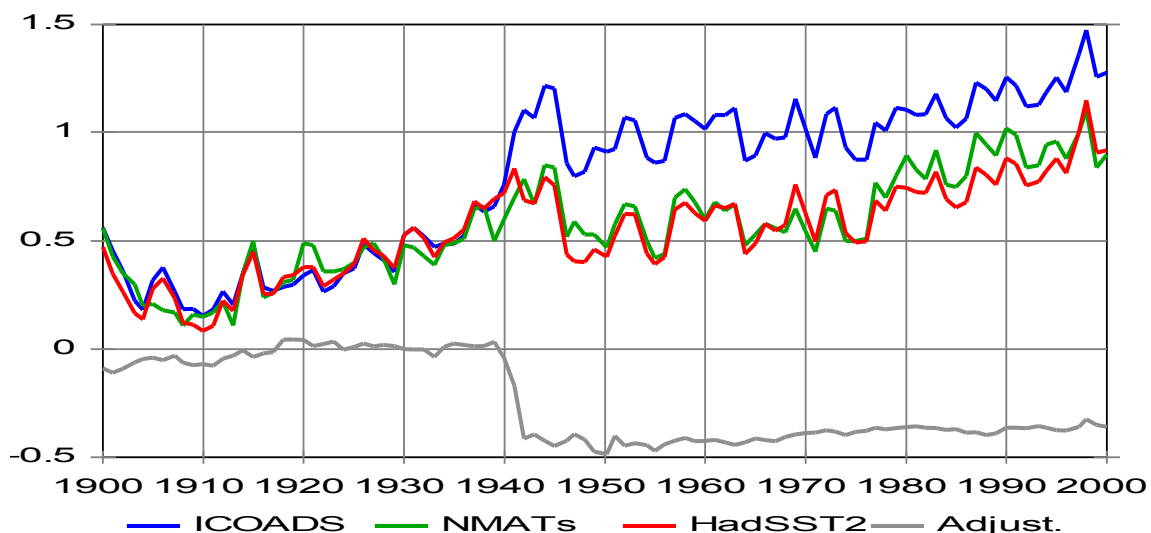
DISCUSSION

We have evaluated the three conditions that must be met before the World War II shift adjustment can be considered valid, and none of them are. Consequently, the adjustment must be considered invalid. What are the impacts?

According to accepted criteria, not very large at all. In 2008 the adjustment was in fact invalidated by Thompson et al. (T08) ⁽¹⁶⁾, who finally identified the 1946 downward shift that terminates the World War II SST spike as artificial, attributing it to an intake-back-to-bucket change after World War II that offset the bucket-to-intake change that caused the upward shift in the SST record earlier in World War II. T08 recognized that “new corrections” were necessary to remove the 1946 shift but did not expect them to alter “estimates of the century long trend in global-mean temperatures”. But removing the 1946 shift removes the permanent shift correction, and this has a very large impact on long-term temperature trends. So how did T08 conclude that it would not?

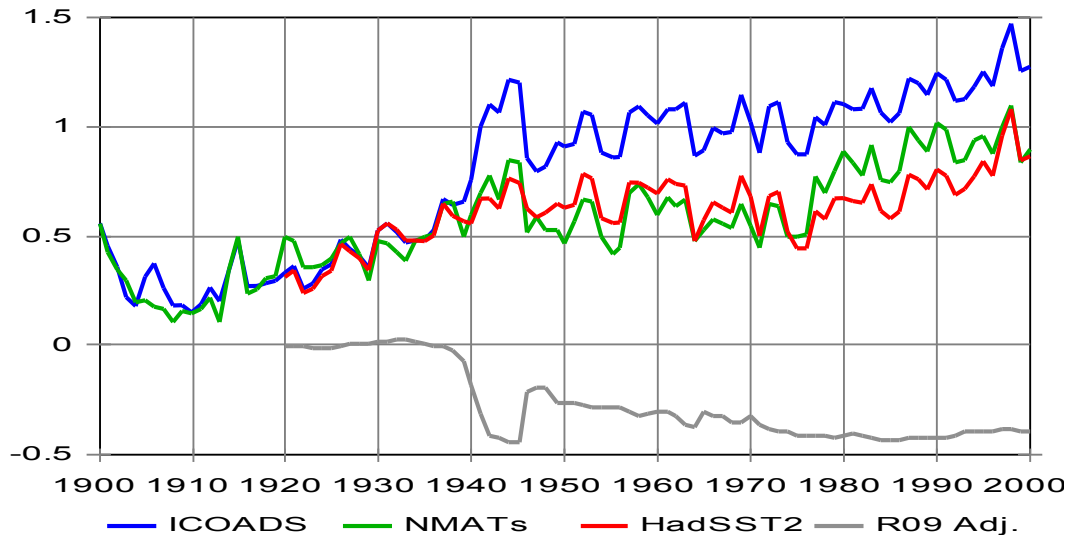
Because SST bias adjustments are based on the assumption that SST trends are representative of surface air temperature trends over the oceans (note that SST series are actually proxy reconstructions). This assumption has to be made or SSTs cannot be used to construct air temperature series, but once it is made the final SST series (in this case HadSST2) must match the air temperature series (the NMAT series) because if it does not it is obviously not representative of ocean air temperatures and therefore by definition invalid. But HadSST2 already matches the NMAT series (Figure 12), meaning that it must be basically correct as it stands, which in its turn means that the permanent shift adjustment must also be basically valid. And if this adjustment is valid and a single bucket-intake change does not explain it, then there must be some other previously-unidentified bias in the SST record that does.

FIGURE 12
SST, NMAT AND HADSST2 SERIES WITH SHIFT ADJUSTMENT (Degrees C, annual means)



As T08 saw it, therefore, we had the final result right but had misidentified the sources of SST bias. The solution, therefore, was to rework the bias adjustments to a) remove the artificial 1946 shift and b) identify the previously-unrecognized bias that was needed to restore the permanent shift adjustment. Revised adjustments to do this have now been proposed by Rayner et al. 2009 (R09) ⁽²¹⁾, and these are summarized in Figure 13 (data from R09 Figure 1).

FIGURE 13
SST, NMAT AND HadSST2 SERIES WITH R09 ADJUSTMENTS (Degrees C, annual means)



R09 do not describe the specific procedures used to develop these adjustments, but they allow for two main sources of bias. First is an assumed increase in intake measurements from 0-10% before WWII to 20-50% after (R09's intake percentages are similar to the K10 estimates in Figure 4). This adjustment is an improvement over the permanent shift adjustment in that it assumes more realistic bucket-intake ratios, but it is still not well-supported by metadata and it will again tend to be too large because R09, like K10, also underestimate prewar intake percentages. It also visibly undercompensates for the 1946 shift, which has an amplitude of 0.4C (Figure 5) but which receives an adjustment of only 0.2C.

Second is an assumed postwar transition from dominantly uninsulated to 100% insulated bucket measurements. There is some evidence that such a transition might have occurred, but it had never previously been thought to have had any significant impact on the SST record. Moreover, if there was such a transition we have no way of estimating how uninsulated-insulated bucket ratios changed with time, if indeed they changed at all, because we have no bucket-type metadata, and because of this lack of metadata we are also unable to make any analytical estimates of uninsulated-insulated bucket bias levels. Finally, the R09 bucket-bias adjustment assumes that the SST record contains a 0.2C warming bias between 1950 and 1980, and there is no evidence for this. The SSTs in fact match the NMATs more closely over this period before the bucket-bias adjustment is applied (Figure 12) than after (Figure 13).

In summary, there is essentially no more support for R09's revised adjustments than there was for the permanent shift adjustment they are designed to replace. But they re-match the SSTs to the NMATs, which makes them conceptually valid.

It can be argued that this approach is unscientific because it implicitly validates any bias adjustments that match the SSTs to the NMATs whether they are valid or not. However, it is the approach that has historically been used, basically because in the absence of metadata there is no other way of doing it, and if SSTs are valid air temperature proxies it will indeed give the right results. But if they are not then all four IPCC Assessment Reports will have based their conclusions on surface air temperature time series that are at least technically incorrect.

Clearly the entire SST edifice – and to some extent the entire “global warming” edifice – stands or falls on the single issue of whether SSTs really are valid air temperature proxies. The next section addresses this issue.

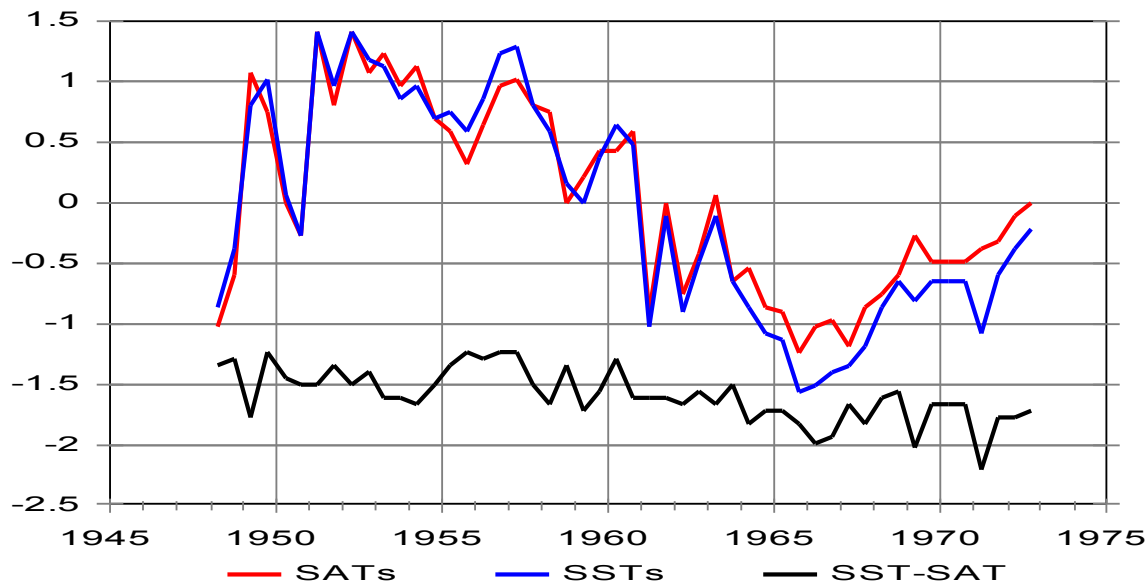
ARE SSTs VALID AIR TEMPERATURE PROXIES?

The US National Resource Council's 2005 Review of the U.S. Climate Change Science Program's Synthesis and Assessment Report contained the following observation: “In lines 107-111, the benefits of sea surface temperature (SST) over night marine air temperature (NMAT) are discussed without saying anything about what the relationship between SST and NMAT is likely to be (e.g., is SST a good proxy for NMAT?)”

In fact it appears that very little has ever been said about what the relationship between SST and NMAT is. To the best of this writer's knowledge there are no statements anywhere in the literature to the effect that “SSTs are valid air temperature proxies, and here is the proof”. All of the available evidence is indirect. Here we will summarize this evidence in a brief historical review.

The earliest study to address the question of whether SSTs are valid air temperature proxies – and apparently the only one that has ever studied it directly – was Cayan (1980)⁽²²⁾. Cayan compared over ten million SST readings with the same number of surface air temperature readings taken in the North Atlantic and North Pacific between 1948 and 1972 and found that “sea surface temperature ... and surface air temperature ... are well related on monthly, seasonal and annual time scales”, which indeed they are. However, he said nothing about how well-related they were over time scales longer than a year, and in fact did not even address the issue. Had he done so he would have found that SSTs and surface air temperatures were not always well-related even over the short 25-year comparison period he used. An example is shown in Figure 14, which reproduces the SST versus surface air temperature comparison for the 10x10 degree North Atlantic square shown in Cayan's Figure 1. The difference plot shows surface air temperatures warming by ~0.5C relative to SSTs between 1957 and 1972, and this is not readily explicable as a bias effect. (There is no known reason why there should be any large biases in either the SST or the NMAT records over this period, and if one believes the R09 bias adjustments shown in Figure 13 then air temperatures should have cooled relative to the SSTs over this period, not warmed.)

FIGURE 14
CAYAN 1980 AIR TEMPERATURE-SST COMPARISONS (Degrees C, annual means)
North Atlantic 10 degree grid square (45N 65W)

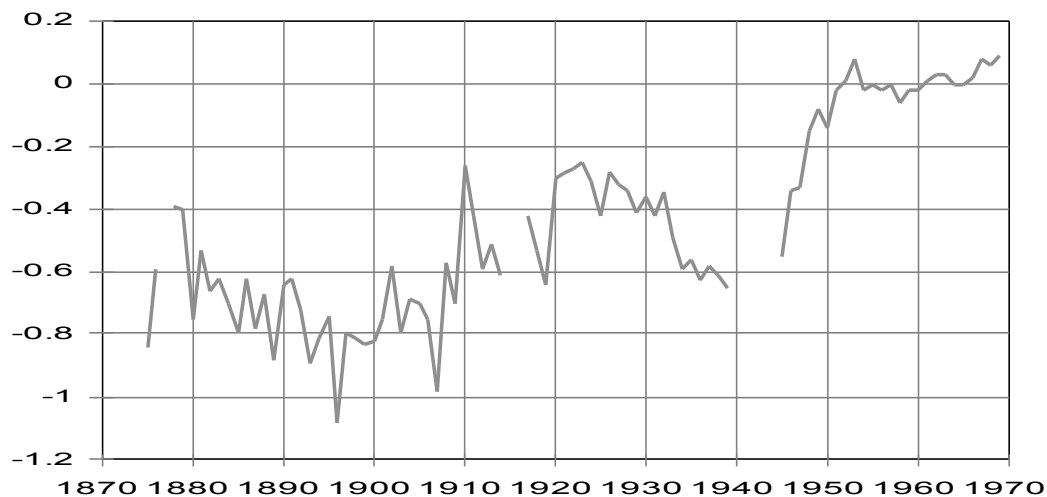


The problem with Cayan's conclusions, however, was that SSTs could be used to construct century-long air temperature time series only if they were valid air temperature proxies over all time scales, not just over time scales of up to a year. It was therefore perhaps inevitable that Cayan's conclusions would be misrepresented. Paltridge and Woodruff (1981) ⁽²³⁾ were the first to do this: "Cayan (1980) has shown that SST variability is generally a good indication of surface air temperature variability on time scales of more than a month". Then in 1984 Barnett ⁽⁷⁾ stated: "beyond the seasonal averaged time scale, air temperature changes over the oceans are closely related to sea surface temperature changes (Cayan, 1980)." In 1990 Servain et al. ⁽²⁴⁾ stated: "Indeed, for data collected from the beginning of the present century, it is generally accepted (Cayan 1980, Paltridge and Woodruff 1981) that these two variables (SSTs and air temperatures) vary in harmony on time scales of more than a month". Thus did Cayan's study help cement in place the concept that SSTs were valid long-term air temperature proxies even though it actually said nothing of the sort.

And once this concept was cemented in place divergences between the SST and the NMAT series became, by definition, a result of biases in the SST record. This was the basis for the 1984 Folland, Parker and Kates adjustments. It was also used by Oort and Maher (1985) ⁽²⁵⁾, who compared SST and marine air temperature readings in the tropics between 1920 and 1975 and found SSTs warming by about 0.6C relative the MATs between 1939 and the late 1950s, similar to the divergence found by Folland, Parker and Kates (Figure 11). Oort and Maher attributed this divergence to a "change from bucket to engine intake measurements", dismissing the possibility that it might be natural because 0.6C of differential warming "impl(ied) a change in the net sea-air exchange of energy on the order of 20 W m⁻², too large to be acceptable." However, more recent work has found that a 20 W m⁻² change is not necessarily too large to be acceptable. According to the IPCC AR4 (Section 3.2.2.3) ⁽²⁶⁾ variations in sea-air heat fluxes "exceed ± 50 W m⁻² for many months during major ENSO events", and according to Yu and Weller (2007) ⁽²⁷⁾ global sea-air heat flux increased naturally by about 10 W m⁻² between 1981 and 2005.

Next came Wright (1986) ⁽²⁸⁾, who compared approximately ten million SST and NMAT readings taken in the North and South Pacific between 1870 and 1969. Figure 15 shows Wright's SST minus NMAT difference plot, reproduced from Wright's Figure 5. The SSTs warm erratically by about 0.8C relative to the NMATs between about 1900 and 1970. Wright also attributed most of this divergence to “changes in the method of measurement of both SST and (air temperature)”, a conclusion clearly based on the assumption that the SSTs should match the NMATs. It was certainly not backed up by metadata (“much of the required information about individual observations is not known (for example, whether SST was measured by bucket or intake)”).

FIGURE 15
WRIGHT SST MINUS NMATS, PACIFIC OCEAN (Degrees C, annual means)



Following Wright came Bottomley et al. (1990) ⁽²⁹⁾ and Folland and Parker (1995) ⁽¹²⁾. These studies also provided no hard evidence that SST-NMAT divergences were a result of biases. They simply assumed that they were and proceeded to use SST-NMAT differences to quantify these biases, with Bottomley et al. using the differences to quantify biases in the NMAT record and Folland and Parker using them to reconstruct uninsulated-insulated bucket ratios (they had no bucket-type metadata) that were then used to quantify biases in the SST record. The fact that no one saw anything wrong with these approaches demonstrates that the SST=NMAT concept had already taken firm hold.

The next publication to address the issue was the 2001 IPCC TAR ⁽³⁰⁾. Here the question of whether SSTs were valid air temperature proxies was addressed indirectly via studies which had reportedly confirmed the Folland and Parker adjustments. Two studies were cited. First, Hanawa et al. (2000) provided “independent confirmation” of these adjustments by comparing air temperatures and SSTs at five coastal stations around Japan, but this was obviously too small a sample to be diagnostic. Second, Folland et al. (2001) forced a climate model using two SST series, one that contained the World War II shift adjustment and one that did not, and compared the output with a “land air” temperature series (the results are summarized in Figure 2.4 of the TAR). Air temperatures forced with the shift-adjusted SSTs matched the land air series and air temperatures forced with the unadjusted SSTs did not, thereby providing “Confirmation that the (Folland and Parker) adjustments are quite realistic globally.” This conclusion is questionable, but the Folland et al. results did show that SSTs should track ocean air temperatures according to climate models, which is arguably the closest anyone has come to proving that SSTs are valid air temperature proxies.

But empirical comparisons continued to show divergences between the SSTs and the NMATs that were not readily explicable as bias effects, and gradually the possibility that they might be natural began to be considered. In 2001 Christy et al. ⁽³¹⁾ found that “the rate of atmospheric warming in the tropics since 1979 is less than the observed warming of the sea surface”, suggesting a possible natural origin. In 2002 Smith and Reynolds ⁽³²⁾ noted that “climate variations such as the North Atlantic Oscillation may also cause changes in the SST–NMAT difference”. In 2003 a Folland et al. Study ⁽³³⁾ compared SSTs and island air temperatures in four regions in the South Pacific and concluded that “trends in ... annual and seasonal marine and island temperatures in the South Pacific often show good agreement but there are differences whose origin may be either poor data *or climatic*.” (author's italics). Current series also show large divergences between the SSTs and the NMATs (Figures 6, 7 and 8) that are not readily attributable to bias effects, as previously noted.

The 2002 Smith and Reynolds study cited above is also noteworthy from a different standpoint. Its goal was to obtain another independent verification of the Folland and Parker bias adjustments by basing bias corrections entirely on SST-NMAT differences, but it differs from other studies in that it treats the the SST=NMAT relationship as an assumption, not a fact, as the following quotes attest: “Specifically, we assume that ... any changes in the magnitudes of the (SST-NMAT differences) were attribut(able) to measurement or instrument changes, and these changes are assumed to affect only SST”. “These assumptions are unlikely to be completely true”. “Our most questionable assumption is that historic changes in SST–NMAT reflect only SST changes”. “We will proceed with these assumptions with the realization that they introduce some uncertainty in our results”.

Since 2003 the SST=NMAT issue has been largely ignored. Rayner et al. (2005) ⁽³⁴⁾, T08 and R09 say nothing about it. The question was briefly revisited by K10, who claimed that: “Anomalies in marine air temperatures (MAT) tend to be correlated with anomalies in SST over long space and time scales” – hardly a ringing endorsement (note the phrase “tend to be”). In addition, no citation was provided, which is surprising given that K10 reference no fewer than 77 previous studies. The only study K10 reference relative to the SST=NMAT relationship is in fact Smith and Reynolds, which as noted above does not accept the relationship as proven either.

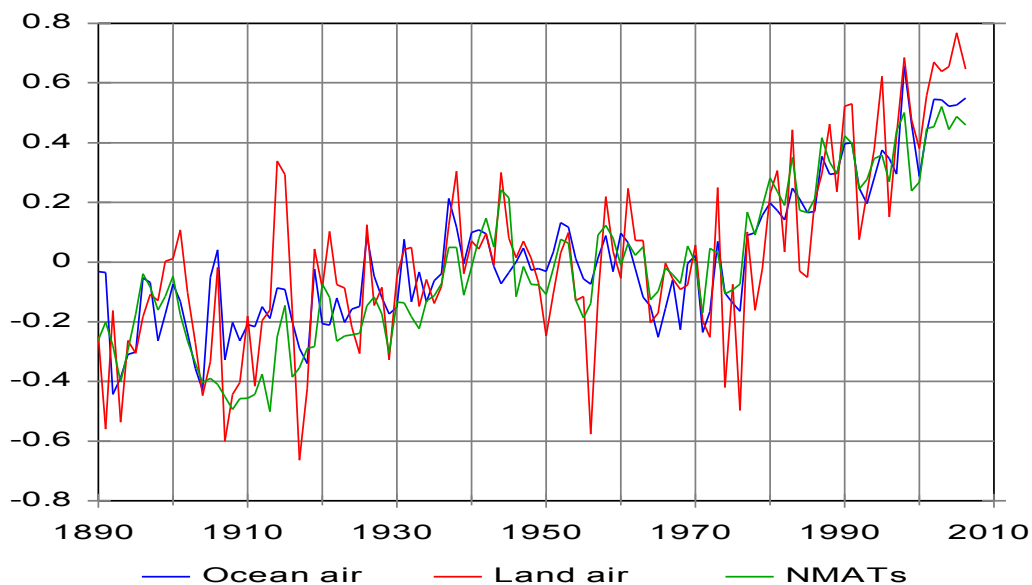
This brief historical review demonstrates that we are unable to prove that SSTs are valid surface air temperature proxies over all time scales. We are also unable to prove that they are not, but that is not the point. Before we can use SSTs to construct air temperature time series we must be certain that SSTs really are valid air temperature proxies, and we are not. Consequently we should not be using SSTs to construct air temperature time series. (Another reason for not using them is that if the ultimate goal of SST bias adjustments is simply to match the SSTs to the NMATs, which to all intents and purposes it now seems to be, we might as well just keep the NMATs and discard the SSTs altogether.)

WHITHER SSTs?

The fundamental problem with the SST record is that while there are reasons to suspect that it may be biased by measurement method changes, there is no objective way of quantifying these biases or even of proving that they exist. This situation is entirely a result of the lack of SST metadata, and since we are unlikely to discover any new metadata at this late stage we can assume that it will not change. The bottom line is therefore that we will *never* be able to quantify instrumental biases over most of the SST record, or indeed prove that there are any. (The exception is after about 1970, when there does appear to be enough metadata to allow biases to be identified and quantified analytically.)

So if we cannot use SSTs to construct ocean air temperature series, how do we construct them? The answer is simple. We use records that measure air temperatures directly. Figure 16 compares three air temperature series – an “ocean air” series based on unadjusted data from 100 stations on islands on the oceans and a “land air” series based on unadjusted data from 497 stations in landlocked areas, both constructed by the writer using GISTEMP data, and the NMAT (MOHMAT) series. The ocean air and land air series show the same amount of warming and the same multidecadal trends – demonstrating, incidentally, that there are no appreciable differences between temperature trends over the land and the oceans – and the NMAT series tracks them quite closely except around 1910. Clearly we can construct representative ocean air temperature series from land station records alone. (We could also upgrade these series significantly by combining land station and NMAT data. It is in fact legitimate to question why this is not already being done.)

FIGURE 16
OCEAN AIR, LAND AIR AND MOHMAT NMAT SERIES (Degrees C, annual means)



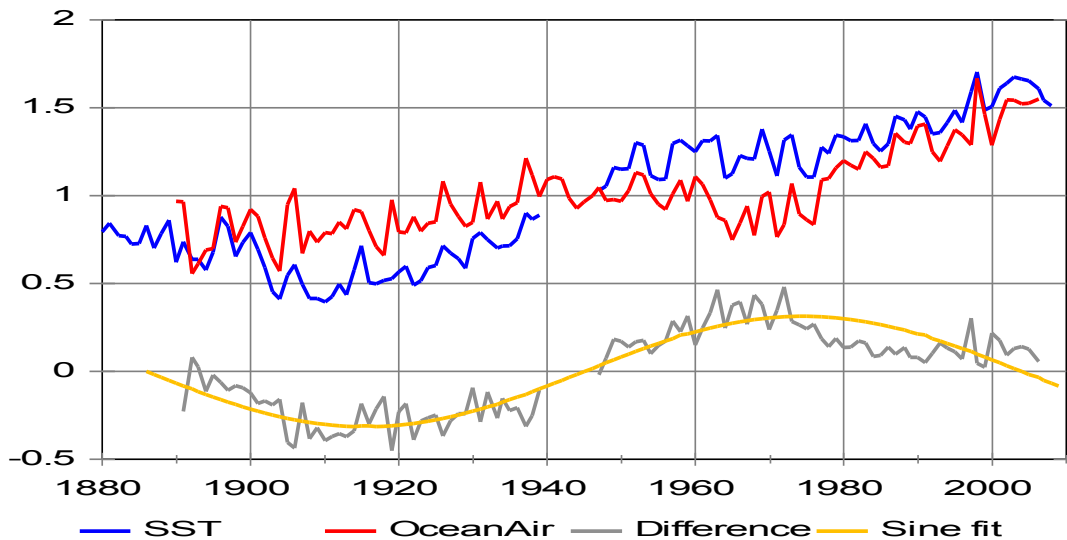
As to what to do with the SST series, we clearly cannot ignore it because the heat content of the oceans greatly exceeds the heat content of the atmosphere. Accordingly, the following suggestions are made:

1. Make no assumptions as to whether SSTs are or are not valid air temperature proxies. Consider the SST series on a stand-alone basis. Treat it independently of surface air temperature series in the same way as surface air temperature series are treated independently of the MSU lower troposphere series.
2. Assume that the unadjusted SST record is not biased by instrumental changes (there is no proof that it is, only supposition).
3. Delete the contaminated World War II data (they are essentially unrecoverable anyway).

When we adopt these suggestions some interesting patterns begin to emerge. These are summarized in three final figures, which are not intended to prove anything but rather to highlight certain possibilities which up until now do not appear to have been considered.

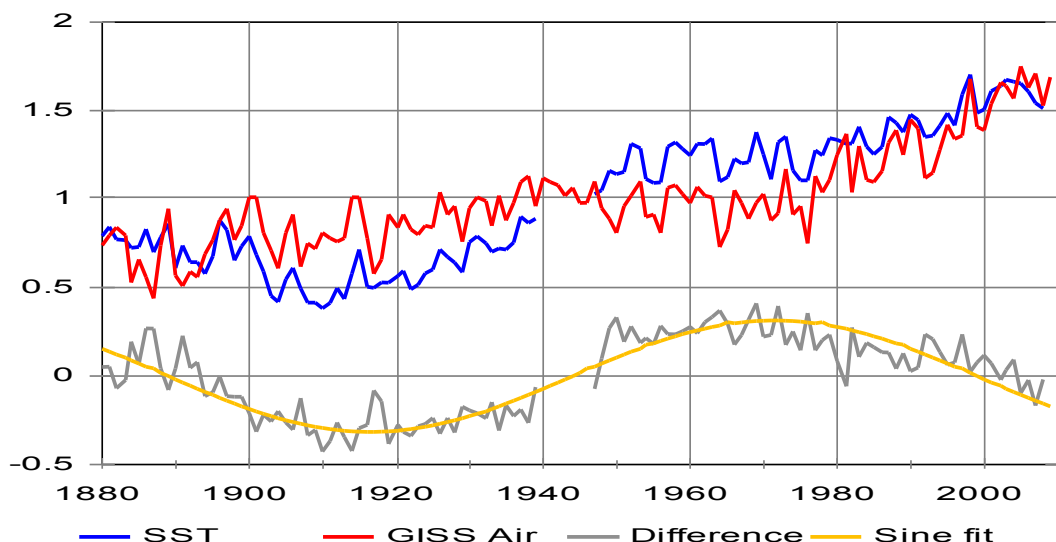
First, it will have been noted in earlier Figures (e.g. Figure 6) that SST-NMAT difference plots tend to exhibit alternating periods of differential warming and cooling. Figure 17 now compares the ICOADS SST series with the “ocean air” series shown in Figure 16. This is a robust comparison because the two series are independent, cover only ocean areas and are unadjusted. Both show about the same amount of overall warming between 1890 and the present, but over the intervening period they oscillate around each other. The oscillations are not completely regular, but as shown by the difference plot they can be fitted quite closely ($R=0.91$ for annual means and $R=0.95$ for ten-year smoothed means) with a sine curve that has an amplitude of 0.63°C and a period of 118 years.

FIGURE 17
ICOADS SST VS. “OCEAN AIR” SERIES (Degrees C, annual means)



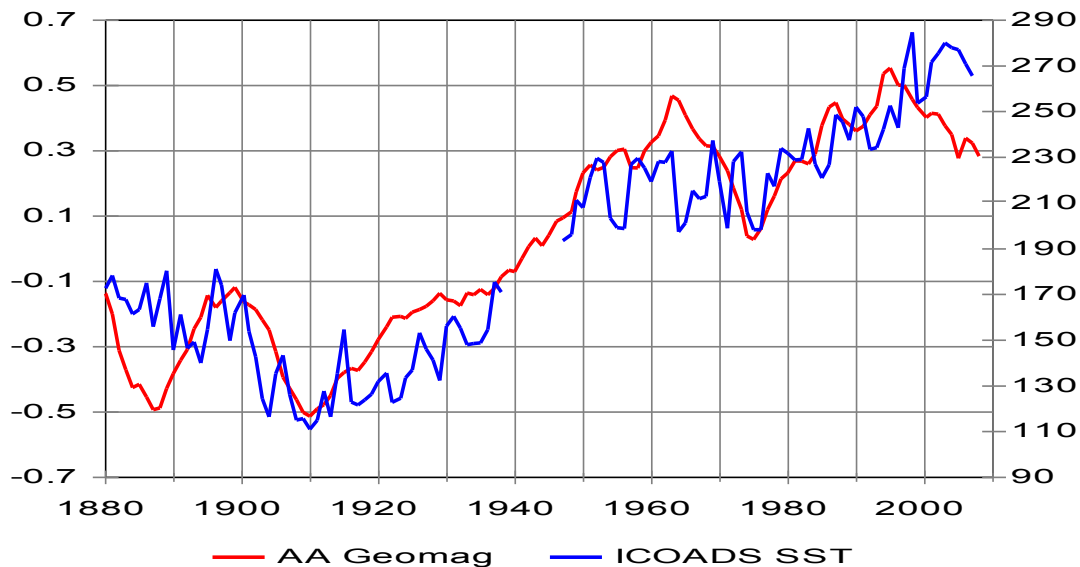
Comparing the ICOADS SST series with other surface air temperature series, such as the GISS global “meteorological station” series ⁽³⁶⁾, the “land air” series shown in Figure 16 and the HadMAT series shown in Figure 9, gives similar results. The comparison with the GISS series is shown in Figure 18.

FIGURE 18
ICOADS SST VS. GISS METEOROLOGICAL STATION SERIES (Degrees C, annual means)



Second, the divergence between the SST and NMAT series between roughly 1920 and 1960 is mostly a result of the strong warming in the SST record over this period (Figure 7, for example, shows 0.8C of warming between 1910 and 1960). Figure 19 now compares the ICOADS SST series with the AA geomagnetic index ⁽⁴⁰⁾, with the AA index shown as smoothed 11-year means and advanced by nine years to simulate a nine-year lag in the SST series relative to the AA index. The 1910-1960 warming in the SST series coincides with a period of rapidly increasing geomagnetic activity. Substantially the same pattern is seen when the SST series is compared with solar irradiance reconstructions.

FIGURE 19
ICOADS SST SERIES VS. AA GEOMAGNETIC INDEX
(Left scale, SSTs, degrees C, annual means, Right scale, AA index, nT)



REFERENCES:

1. http://jisao.washington.edu/data/global_sstanomts/
2. <http://www.cru.uea.ac.uk/cru/data/temperature/#datter>
3. <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0450%281963%29002%3C0417%3AASOTQO%3E2.0.CO%3B2>
4. http://books.google.com.mx/books?id=NEuXtdcroDMC&printsec=frontcover&dq=roll+1965+physics+marine+atmosphere&source=bl&ots=iGce1jgRzn&sig=dbiU15fxaaumb7TwKAqG7yXyvEY&hl=en&ei=9f0pTMbbGMn6nAebkZihAQ&sa=X&oi=book_result&ct=result&resnum=4&ved=0CCKQ6AEwAw#v=onepage&q=roll%201965%20physics%20marine%20atmosphere&f=false
5. <http://www.springerlink.com/content/7812261801xgnu00/>
6. <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0450%281978%29017%3C0374%3ACOOOSS%3E2.0.CO%3B2>
7. <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0493%281978%29106%3C1353%3AEVOSAT%3E2.0.CO%3B2>
8. http://eprints.soton.ac.uk/69540/1/KentTruscottTaylorHopkins_mmroa26.pdf
9. <http://journals.ametsoc.org/doi/full/10.1175/JTECH1845.1>

10. <http://www.realclimate.org/index.php/archives/2008/06/of-buckets-and-blogs/>
11. <http://wires.wiley.com/WileyCDA/WiresArticle/wisId-WCC55.html>
12. ftp://podaac.jpl.nasa.gov/pub/sea_surface_temperature/buoy/gostaplus/hdf/Document/papers/1-crrt/1-CRRT.HTM
13. <http://journals.ametsoc.org/doi/pdf/10.1175/JCLI3637.1>
14. <http://docs.lib.noaa.gov/rescue/mwr/054/mwr-054-06-0241.pdf>
15. http://books.google.com.mx/books?id=ECMrAAAAYAAJ&pg=PA518&lpg=PA518&dq=Stupart,+F.,+Patterson,+J.+and+Grayson-Smith,+H.,+1929,+Ocean+surface+water+temperatures-methods+of+measuring+and+preliminary+results.&source=bl&ots=tJ957KRbWr&sig=OrQeYKsLIWSzbP5Go8760OO0Jd8&hl=en&ei=EqqDTQquOYjSsAPQ6sn2Bw&sa=X&oi=book_result&ct=result&resnum=1&ved=0CBIQ6AEwAA#v=onepage&q&f=true
16. http://www.atmos.colostate.edu/ao/ThompsonPapers/Thompson_etal_Nature2008.pdf
17. <http://climexp.knmi.nl/start.cgi?someone@somewhere>
18. <http://gcmd.nasa.gov/KeywordSearch/Metadata.do?Portal=GCMD&KeywordPath=&NumericId=19724&MetadataView=Data&MetadataType=0&lbnode=mdlb2>
19. <http://www.ncdc.noaa.gov/bams-state-of-the-climate/2009-time-series/?ts=mat>
20. <http://www.nature.com/nature/journal/v310/n5979/abs/310670a0.html>
21. https://abstracts.congrex.com/scripts/jmevent/abstracts/FCXNL-09A02a-1662927-1-Rayneretal_OceanObs09_draft4.pdf
22. <http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281980%29108%3C1293%3ALSRBSS%3E2.0.CO%3B2>
23. <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0493%281981%29109%3C2427%3ACIGSTF%3E2.0.CO%3B2>
24. http://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_7/b_fdi_59-60/010026217.pdf
25. http://books.google.com.mx/books?hl=en&lr=&id=vQCFsBY0QmkC&oi=fnd&pg=PA183&dq=paltridge+woodruff+sst+1981&ots=1gtxMvgje-&sig=MfoiOOUzRCsIKZPOAC6PLEz_wuM#v=onepage&q&f=true
26. http://www.ipcc.ch/publications_and_data/ar4/wg1/en/ch3s3-2-2-3.html
27. <http://journals.ametsoc.org/doi/pdf/10.1175/BAMS-88-4-527>
28. <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0493%281986%29114%3C1028%3APITUOS%3E2.0.CO%3B2>
29. <http://iridl.ldeo.columbia.edu/descriptions/.GOSTA.html>
30. <http://www.ipcc.ch/ipccreports/tar/wg1/055.htm>
31. <http://www.agu.org/pubs/crossref/2001/2000GL011167.shtml>
32. <http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%282002%29015%3C0073%3ABCFHSS%3E2.0.CO%3B2>
33. <http://journals.ametsoc.org/doi/pdf/10.1175/1520-0442%282003%29016%3C2859%3ATAVISP%3E2.0.CO%3B2>
34. http://camels.metoffice.gov.uk/hadsst2/rayner_etal_2005.pdf
35. <http://data.giss.nasa.gov/gistemp/taledata/GLB.Ts.txt>