A submission to the GWPF temperature data review.

This is a submission from Nick Stokes, Melbourne, Australia. I am a now retired CSIRO research scientist in applied mathematics, though I still hold a post-retirement fellowship. CSIRO had no involvement in this submission. I maintain a blog on climate matters - Moyhu. I write there a lot on temperature measurements. I maintain my own least-squares based global index, TempLS, and for three years have been publishing monthly results, usually in advance of the majors. For the land component, I use GHCN V3 unadjusted, although as I shall show, adjustment makes little difference.

The structure of this submission is thus. I will briefly comment on the terms of reference, then expand a little more on the datasets and their use in indices. Then I will include some blog posts which deal in detail with some of the issues.

Terms of reference

I refer to this GWPF page, accessed 27 April.

1. Are there aspects of surface temperature measurement procedures that potentially impair data quality or introduce bias and need to be critically re-examined?

This is a puzzling remit. Is the panel going to make recommendations about how past measurements should have been made? Or undertake its own independent review of current practice? Without understanding the purpose, I can't comment further.

2. How widespread is the practice of adjusting original temperature records? What fraction of modern temperature data, as presented by HadCRUT/GISS/NOAA/BEST, are actual original measurements, and what fraction are subject to adjustments?

"How widespread is the practice..." is rather unscientific phrasing. People use station data to calculate continuum estimates for the globe and regions. That is basically spatial integration, and to do it, estimates of temperature in local sub-regions are necessary. These are derived from, but not identical with, the raw measures. The estimated value may well be expressed as an adjustment.

HadCRUT/GISS/NOAA/BEST are quite explicit about their adjustment procedures. GISS, for example, uses GHCN adjusted. They publish their code[15]. There is no mystery there. HadCRUT does little adjustment of its own, but uses adjusted data from some suppliers. GHCN adjustment procedures are extensively documented - see reference list[1-12], code[12].

3. Are warming and cooling adjustments equally prevalent?

No. Why should they be? They are made to identify and repair bias. There is no law that that has to be equally distributed. No time scale is given over which the warm/cool is to be measured. Warming usually refers to trend. Below is a plot (from here) of the effect of GHCN adjustment on a global index (TempLS) trend to present, with the x-axis marking the beginning of the period.
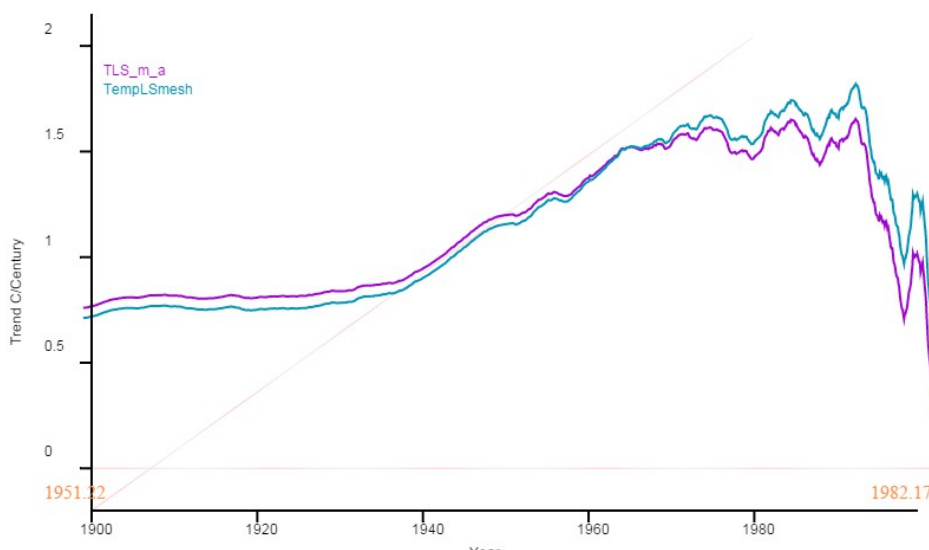
The purple curve is adjusted, blue not. In the more recent period, the adjusted trend is lower; there is a changeover in the 1960's and the longer trends are warmer after adjustment. The differences are small. This reflects that fact that the land component is a small part of global. SST is itself adjusted, but there historically the adjustments have a cooling effect.

4. Are there any regions of the world where modifications appear to account for most or all of the apparent warming of recent decades?

This is a question of great apparent interest to some bloggers, but I can't see that it has any scientific importance. Regional averages attract little attention, and variations in adjustment effect wash out in the global average. Put another way, the popular sport of finding individual warming adjustments should be balanced, if anyone bothered, with the cooling regions that are needed if the average overall is small.

5. Are the adjustment procedures clearly documented, objective, reproducible and scientifically defensible? How much statistical uncertainty is introduced with each step in homogeneity adjustments and smoothing?

Those who like to demagogue these things pay scant attention to the serious scientific work of people like Menne and Williams. Yes, they document and justify their procedures[1-4 *et seq*]. Extensive references are appended. As to statistical uncertainty, again it would have been useful if the remit has specified uncertainty of what.

Global averages are averages of millions of individual readings. Unbiased noise is very heavily damped in the result. The big source of error is bias, which is not damped in the same way. The essential purpose of homogenisation is to identify and minimise this bias. The tradeoff vs added unbiased noise is advantageous, because of that damping.

General remarks

Dr Peiser echoed the announcement of this project by Mr Booker, but sanitised his headline. That headline in the Telegraph was:
"Top Scientists Start To Examine Fiddled Global Warming Figures"
And this is the populist theme, that somehow the record has been altered.

But it hasn't. The original records are actually held by the national Met offices. GHCN, maintained by NOAA, publishes GHCN Daily, which are unaltered Met records of daily max and min, and their average. GHCN monthly has an unadjusted file, which is the simple monthly average of those. They also provide a clearly marked adjusted version. I have seen endless "exposés" were some old record is found and shown to be different to GHCN adjusted. Of course. But it is rarely different to current GHCN unadjusted.

I was a little surprised to see GWPF promote a global temperature review with a graph of USHCN adjustments to CONUS. Contrarian blogs do rarely distinguish, but I hope the panel will. But it is an instructive case. The USHCN adjustment is dominated by TOBS - time of observation. This relates to the process whereby daily min/max are assembled as a month average. In fact, they are readings taken at a specific time in a 24hr cycle, and don't necessarily align with days. There is a considerable bias associated with when the readings are taken. This doesn't matter too much, unless the time of observation drifts. The US has, unusually, largely volunteer observers. That time did drift, and that is documented. And we now have extensive hourly and better records, which give an excellent estimate of the effect of changing times of observation. In those circumstances, adjustment is obligatory.

I'll now append extracts from a series of blog posts on individual topics, with calculated results. First is a series of three on the TOBS issue. One sets out the argument; the second shows a distribution of effects and hsows the documented drift, and the third quantifies in a simple way the bias at a particular location.

I'll then include pointers to posts which show the effects of adjustments in various ways. One is a Google Maps app which allows you to show with colored markers the distribution of GHCN station trend changes due to adjustment. Another is an interface to an excellent set of station docs provides by NOAA, giving the fine details of individual stations. These active aspects require you to link to the Web plots.

The final group are studies of the actual effect of adjustment. One post shows the unadjusted TempLS indices against corresponding adjusted and other indices. It shows both the annual temperatures, and the trend to present (as above). Another shows the breakdown of adjustment effect by continent, and by rural/urban/airport classifications. Again, for these active features to work you need to link to the Web page.

For each blog post I'll add a little yellow box with a note putting it in context, and a link.

TOBS nailed.

Link to original post
This post just sets out the argument for TOBS adjustment. It isn't "altering the record". It is revisiting with modern knowledge the processing needed to get a monthly average from the observations. Details are in the later posts.

It responds to claims that "original data should not be altered" follows my earlier analytic justifications for TOBS (here

and here). See Zeke Hausfather for context, and Victor Venema for a much fuller explanation of min/max thermometers and TOBS.

- The min/max data that you see in a record is not (usually) original data of daily min/max. It is typically a record of the location of min/max markers on a thermometer at a specific time of day (when it was then reset).
- An assumption must then be made to connect that with records of specific days. In the old style, you might assume that a max marker at 5pm Tuesday (example) was the daily max for Tuesday. If it was at 9am, you'd assume it was the max for Monday (and at some time in between, you'd have to switch).
- Repeat, this is an assumption. It is not original data. And it won't always be right. Many of those 5pm Tuesday readings would have been set the previous Monday. That would arise from a warm afternoon when 5pm, not the max for Monday, was warmer than all of Tuesday to 5pm.
- This is double counting, and 5pm creates a warm bias. Warm afternoons can get counted twice. Cold mornings don't.
- Repeating again, an assumption was made and is inevitable. It creates a bias. People raised objections about how the bias can't be measured exactly. I emphasised here that there was a huge amount of data to base an estimate on; that the analysis was straightforward. Oh no, they say, how do you know that people actually read when they said they did (answer - see DeGaetano in that link). Etc. But anyway, the key thing is there is a bias, and it's a scientific duty to estimate and allow for its effect. The objectors want to say it is zero. That's an estimate, baseless and bad. We can do much better.
- The original data is not data about daily temperatures. To get that requires interpretation. And you have to do it right. Laziness won't wash. We can do better. Over the years, NOAA has done better. And yes, for reasons explained in link above, that had a warming effect.

The necessity of TOBS

Link to original
This post shows how observation times in the US have changed over years, and with histograms what effect that will have on various stations. This last is due to Jerry Brennan, writing on the John Daly site.

Recently there has been more said on USHCN adjustments. My recent post on a Steven Goddard plot posted at WUWT produced a response, which conceded the plot was wrong, but WUWT said,
" The one thing common to all of it though is that it cools the past, and many people don't see that as a justifiable or even an honest adjustment."

I've previously written in defence of TOBS, as an adjustment which is not only justifiable but necessary. The information requiring it is staring analysts in the face, and they would be negligent to ignore it. I showed a hourly analysis at Ft Collins, Colo, of the effect of the TOBS bias.

But recently, Steven Mosher reminded that the John Daly site had posted a much more comprehensive survey, by Jerry Brennan in November 2005. The analysis is here, and his summary datafile, which I will use, is here (text, 29kb), which also identifies the stations.

Below is a histogram of the effect of changing TOB from 5pm to 9am for each of the 190 stations considered by Jerry, and of subsequent changes to midnight (standard). There is also a table from the original paper by Karl et el, 1986, which showed that over the years in the US, about 30% of stations made such a change to 1986. Many more stations would have changed to effectively midnight reporting when MMTS came in. The mean effect of the change is 0.66°C cooling. It is no surprise that USHCN adjustments have the effect of "cooling the past".

In the original Karl et el, 1986 paper, there is the following table showing what changes had been made to observing times:

TABLE 1. Percent of A.M. (0100 to 1100 LST), P.M. (1200 to 2100 LST), and MD (2200 to 0000 LST) cooperative weather station in the United States.

| Year	a.m. (%)	p.m. (%)	MD (%)
1931	14	79	7
1941	15	76	9
1951	19	65	16
1965	25	64	11
1975	34	55	11
1984	42	47	11

[Vose et al 2005](#) give this updated plot, showing that the change has gone well towards completion:
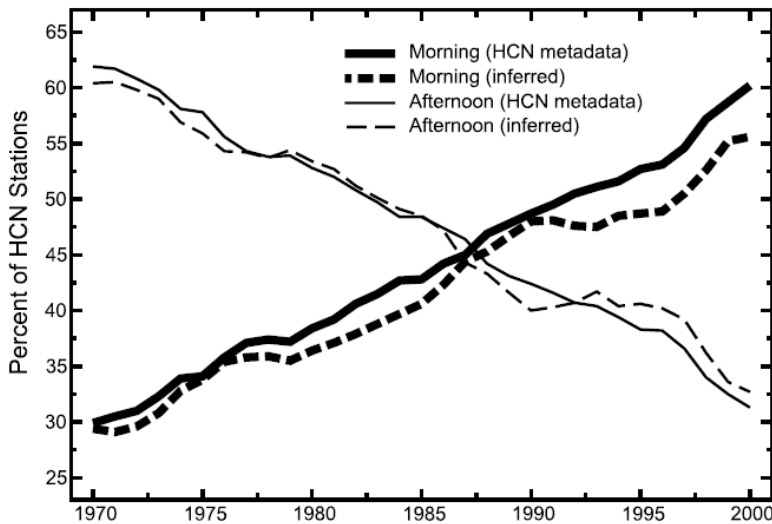
Figure 1. Percent of HCN stations with morning and afternoon observation times according to HCN metadata and metadata inferred using the method of *DeGaetano* [1999].

Because evening TOB has a warm bias, through double counting very warm afternoons, the change to 9am has a cooling effect. Here is a histogram of the putative effects for the 190 stations of changing from 5pm observing time to 9am. Positive effect indicates (in °C) the bias that has to be subtracted from temperatures before the change. The standard setting is midnight, which some stations already observed, and which would become the setting after MMTS conversion. On the right is the required adjustment (in °C) for changing from 5pm to midnight:
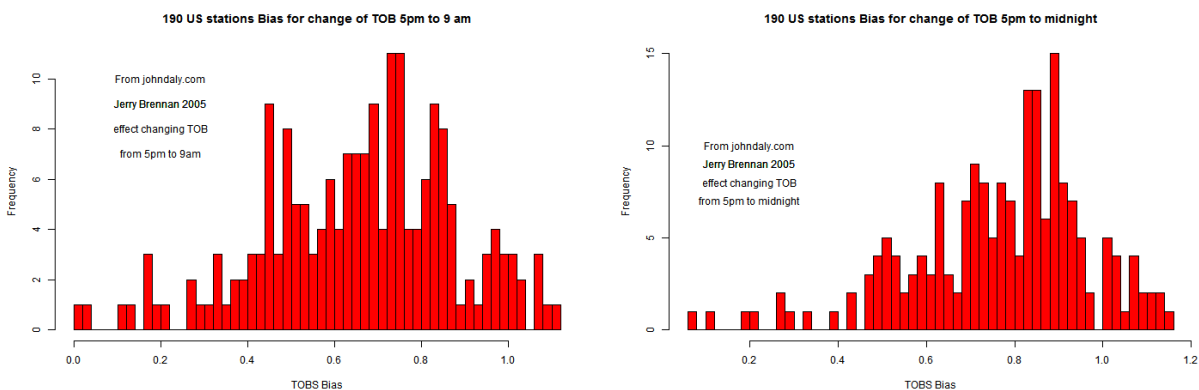
These are large changes, which don't of course apply in full to all stations. A small number were already using midnight. The rest will mostly need some kind of TOBS adjustment, which will "cool the past". But on the evidence, there's no choice.

TOBS pictured

[Link](#)
This shows directly how monthly averaging of various TOB simulations on hourly data gives a reproducible bias.

This is a version of my [TOBS nailed](#) post, with graphics. The numbers come from my [first post](#) in the series, which took three years of hourly data from Boulder, Colorado, and looked at the effect of TOBS (time of observation) measures. That post is the place to look for detail on how it works. A post with much more data is [here](#).

For now, I want to follow the recent post in relating TOBS to fundamentals. What is our measure of average temperature over a period? Sometimes people strenuously urge that the usual TAVG, the average of daily recorded min and max, should be replaced by a proper integral over the day. And they would be right, if we had the historic data. But we don't.

What we do have are records of min and max as recorded daily (at various times of day) by min/max thermometers. These give not the actual daily min/max, but the min/max in the preceding 24 hours (with regular resetting). So they are, averaged over time, a reasonable measure of average temperature, but a measure that depends on the time of observation.

Let me show that with a plot of the three years of Boulder data. I have taken the mean of the hourly data, and compared with the measure that a notional observer would report from reading a min/max every day at at 2AM, or at 5AM and so until 11 pm. I show the 365 day centered running mean that you would get by each of these schemes. The running mean removes the seasonal cycle. The legend shows the colors, with a link to the respective curves. Left axis °F, right in °C. x-axis days after 31 Dec 2008.

Running annual, hourly and TAVG temperature measures, Boulder 2009-2011
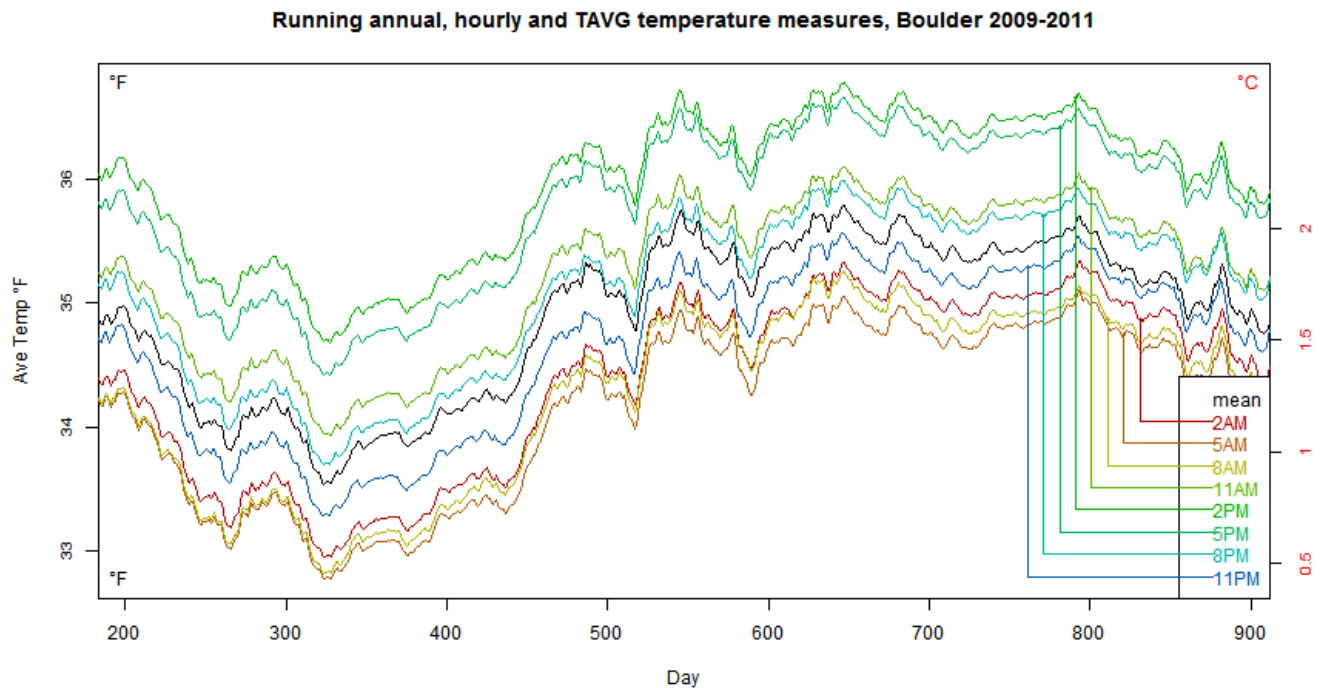
So the various TAVG curves are reasonable measures, in that they track the black mean curve with a roughly constant offset. But the offsets are very dependent on time of obs.

If you stick with one such measure, the offset does not matter much. Its effect would go away on taking anomalies. But if you switch between measures (change TOBS), the effect can be large.

TOBS adjustment is effectively calibrating this measure, relative to a reference. If you change measures, you have to recalibrate.

When we refer to "raw" or "unadjusted" monthly data, it should be remembered that it is not just the original readings. It incorporates an averaging procedure. The outcome of that depends on the time of observation. If that changes, then it's a different measure, as much as if you changed to a differently calibrated thermometer.

Below the fold, I'll show some plots of monthly averages, and a difference plot that may make the stability of the TOBS dependence clearer.

Here now is a version of the above plot shown as differences from the hourly mean. You can see that the curves generally remain in order as TOBS varies, but the spacing is not constant. TOBS adjustment will not be precise year-to-year. But fortunately for most applications, a number of years will be averaged. And there is a lot more hourly data available near most places, so an accurate mean adjustment can be derived.

Diff of TAVG and hourly mean, running annual
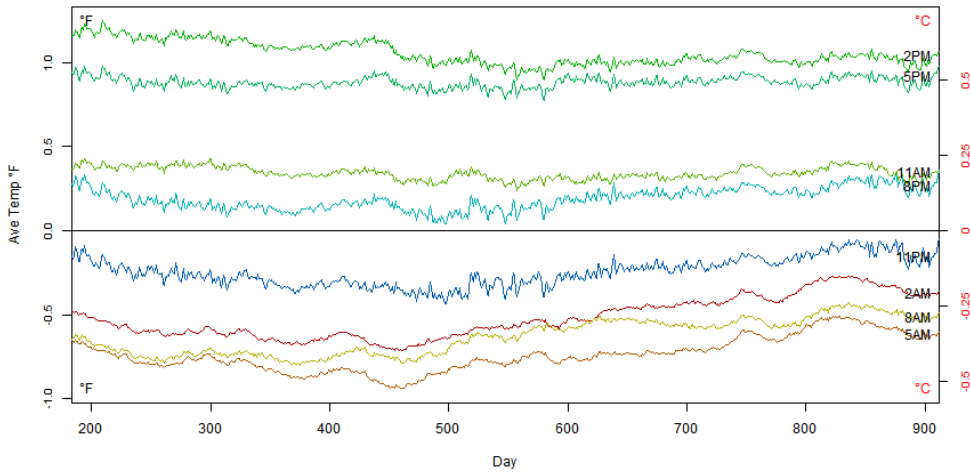
The "coolest" measure is at 5am, when it often happens that a cold minimum is split and attributed to two separate days. 8am is similar but 11am is already on the warm side of neutral, and 2pm is the time most likely to split and double count warm afternoons, making a warm bias. 5pm is a bit less so, and 11 pm has a slight cool bias relative to hourly.

Google Maps and GHCN adjustments

Link to original post
This gadget allows you to display markers on a Google Map (with usual facilities). You can color the markers according to various ranges of adjustment effect. Each marker has a balloon of info including effect. remember, it just shows an image here.

A fortnight ago I posted a Google Maps gadget for viewing GHCN stations colored according to the effect on them of GHCN adjustments. I've been doing some improvements, and rewriting the code in the process. This simplifies the logic, and I'm hoping to produce a generic application to operate on any supplied data.

The plot is below. And below that, some details about the usage logic. The field Trend_Adj is the trend difference over whole of life made by adjustment, in °C/cen. It is set to NaN for stations with less than 360 months of adjusted data in total (maybe with gaps).
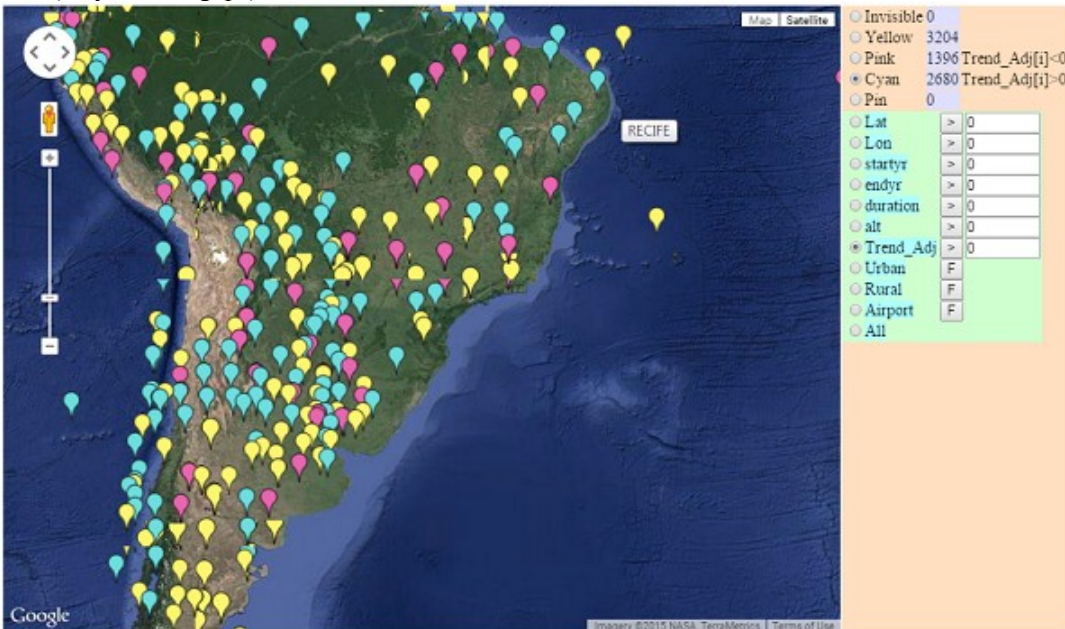
[See the post for operating details]

NOAA GHCN station portal

Link to original post
NOAA has an excellent series of one page summaries of the data for each of their 7280 stations. It shows the months of data present or missing, along with graphs before and after adjustment (with trend) and some histograms. Normally you have to know the code number to access, but this post gives a list of links by name.

NOAA has a rather unpublicized [collection](#) of visualisations of station data. I'll show Honolulu below the jump. It gives a visual summary of even monthly data, contrasts adjusted and unadjusted, including trends. You can enlarge in a new window.

Unfortunately, access seems to be only via an ftp directory in which the filenames are just number codes. Further it is broken into subdirectories which take a long time to load (just the filenames). So I thought I'd develop a portal. The one below is effectively just an extract of the inventory file, but with each station name linked directly to its NOAA information.

Here is the station level information for Honolulu. It has been talked about lately because there has been recent decline, and in recent months the adjusted value has not been included. But it's actually of interest because the unadjusted data over the long term rises steeply, and the adjustment brings the trend down.
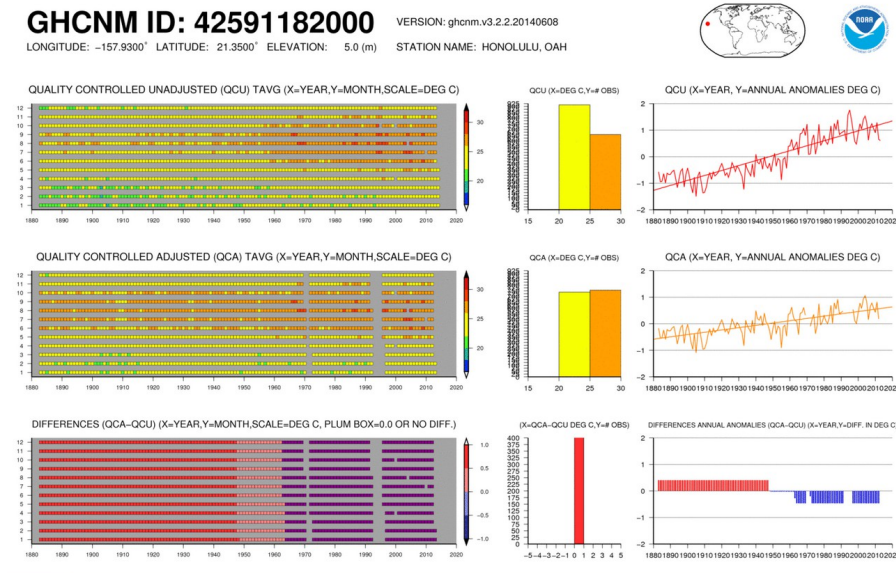
Now here is the list. As well as the station name and country, it shows the number of years of data, and the most recent year in which there is data. You can do text search (Ctrl-F) within the frame.

Station	Yrs	End Data	Country
SKIKDA	63	2014	ALGERIA
ANNABA	52	2014	ALGERIA
DAR-EL-BEIDA	148	2014	ALGERIA
FT. NATIONAL	70	1961	ALGERIA
CAP CARBON	10	1960	ALGERIA
BEJAIA	46	2014	ALGERIA

Homogenisation makes little difference to global average

[Link to original post](#)
This post does a direct comparison of the effect of adjustment on time series and trend. In the active version, you can choose among various datasets.

There has been much blog chatter (reviewed [here](#)) recently about homogenisation of temperature, and adjustment in general. A few individual stations have been picked out and pored over. But homogenisation is a general effort to reduce bias prior to computing a global average, and the logical place to look for its effect is in that average.

This was something some bloggers were interested in doing back in 2010. Zeke writes on that [here](#). His own investigation was mainly US. At that time, I started the [TempLS code](#), and I've been using it for monthly reporting for over three years. It uses by default unadjusted GHCN land data, with ERSST for ocean. There general experience, noted back in 2010, was that it made little difference. TempLS compares well with the major indices.

In this post, I'll try to quantify that more, using the current improved graphics. I can compare directly the variants of TempLS with and without adjustment, with an active map below. The results are a little surprising, but the end effect is still small. A typical result is for TempLS mesh, where the trend 1910-2014 is 0.711°C/Cen, or 0.759°C/Cen after

adjustment. But that is actually a high point of adjustment effect, and over more recent periods, adjustment actually has a cooling effect. Claims that AGW is a creature of adjustments are way off.

My previous post reviewed the analysis of station trends. They tend to be dominated by relatively short periods, even with a cut-off. It is fairly easy for a short term to produce a high trend, but it's the longer terms that contribute most to the average. I should mention a 2012 benchmarking paper by Victor Venema and many others.

The map below is a variant of the one on the latest data page, maintained to date. Operation details are here.
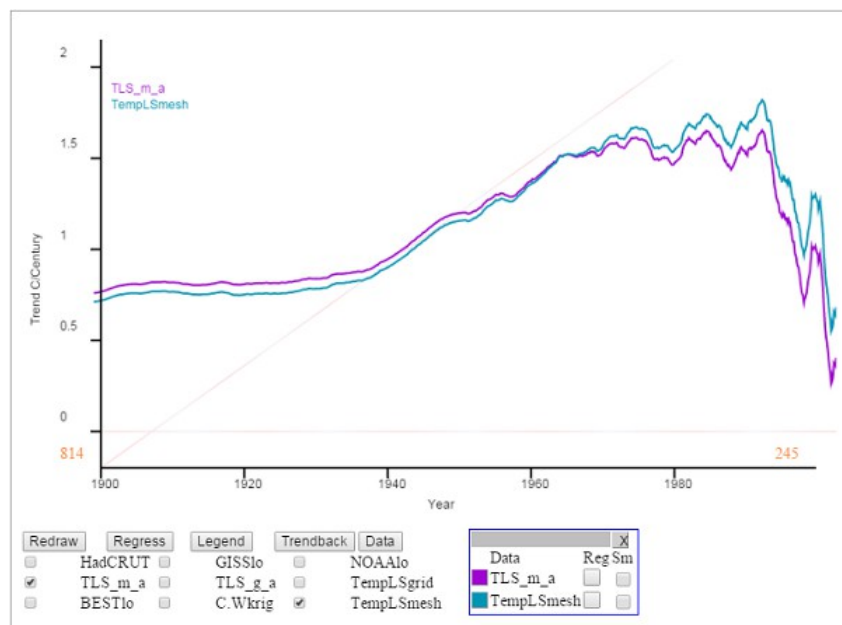
Looking just at the TempLS mesh curves, from about 1960 back to 1900, the adjusted trend is higher. That is the "cooling the past", but it isn't much. Max about 0.04°C/Cen. After about 1970, the unadjusted trend is higher. If you switch to timeseries most (click Trendback), it seems that adjustment has relatively cooled the global temp in just the last few years, which affects the short term trends. But again, it isn't much. It looks a lot in trend, but short term trends are volatile. The grid version of TempLS doesn't really show this. It could be an Arctic effect.

Another way of looking at it is that the difference between adjusted and unadjusted is about the same as that between GISS and HADCRUT.

Breakdown of effects of GHCN adjustments.

Link to original post
This post gives a finer subdivision among various subgroups - regions, urban/rural, airports. Some interesting results.

In a previous post, I showed how running TempLS with unadjusted or adjusted GHCN station data made a small difference to the final global average. That is useful to know, because some say that AGW is a creature of GHCN adjustments. This is certainly not true.

That analysis can be extended. I can trace the contribution to this small adjustment effect from various categories of stations - by continent, rural status, or airport status. And as before, it shows the difference by year, or effect on trend (to present time) by year.

A useful addition to the present post will be some trends calculated for land only in a NOAA paper by Lawrimore et al, 2011, Table 4 (h/t Victor Venema):

|	v3 Uncorr	v3 Corr
1880–2010	0.61°C/Century	0.79°C/Century
1901–2010	0.70°C/Century	0.91°C/Century
1951–2010	0.16°C/Decade	0.18°C/Decade
1981–2010	0.27°C/Decade	0.27°C/Decade

Note the change of units (in the original). I got an approx land only difference of 0.17°C/Cen for 1900-2014 by dividing the global value by the land area fraction. That was pretty rough; the value in this post is more accurate, and agrees well.

Global with urban/rural breakdown

Here is the plot. All have a slight smoothing over 3 years. Note that all these plots can be generated on the zoomable active plotter, along with the associated trend graphs.
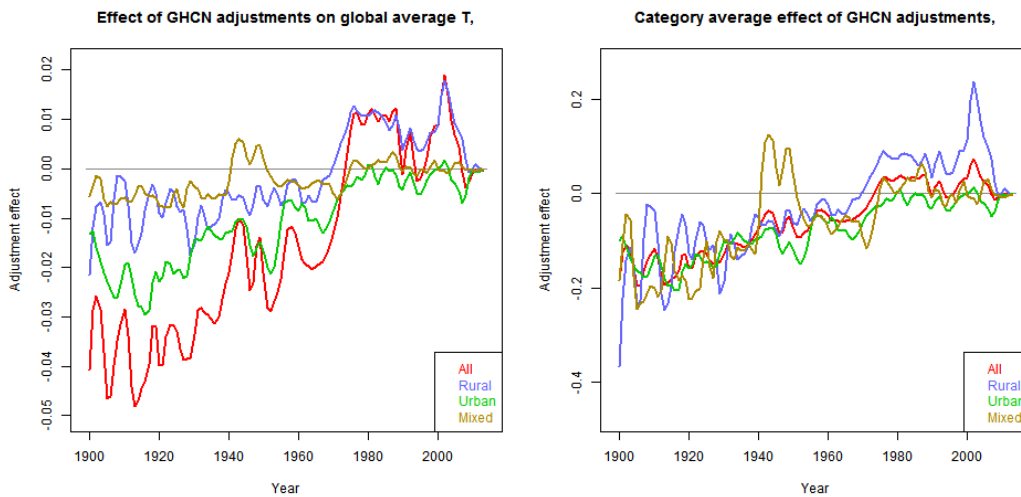
So you can see that the total global contribution is negative to about 1970, then positive. This has the effect that the contribution to trends from about 1960 onward is actually cooling. These are the trends that are vigorously debated as possible evidence of AGW. They are not enhanced by adjustment - on the contrary.

Before 1970 the adjustments "cool the past", by up to 0.05°C. However, on a land basis, that is up to 0.2°C. That is representative of the maximum average adjustment.

A virtue of the left style is that it is additive. So urban adjustments make up the larger part, but not by all that much.

There is an interesting sharp rise to about 1975, and then a decade plateau. We'll see more of it. It seems here associated with rural. My theory is that it is an airport effect, and may be associated with the move of some rural classified stations to airports.
[There are various other breakdowns in the original post.]

Appendix

Here are some more URLs from Moyhu concerning adjustments:
http://moyhu.blogspot.com/2015/02/homogenisation-makes-little-difference.html
More on the rationale for adjustment, even where it has little effect.
http://www.moyhu.blogspot.com.au/2012/02/study-of-ghcn-v3-homogeneity.html
An earlier study of histograms of trend effect, with a Google Maps app.
http://moyhu.blogspot.com.au/2014/05/ushcn-adjustments-averages-anomalies.html
Exposing a blog fallacy in averaging re USHCN.
http://moyhu.blogspot.com/2014/07/ushcn-adjustments-plotted-for-usa-and.html
A complete graphing of USHCN adjustments affecting the various states.
http://moyhu.blogspot.com/2012/10/a-necessary-adjustment-time-of.html
An earlier version of one of the included posts
http://moyhu.blogspot.com/2014/08/amberley-bom-and-wuwt.html
Some Australian adjustments

And here, in order, are the URLs for the posts cited above
http://www.moyhu.blogspot.com.au/2014/06/tobs-nailed.html
http://www.moyhu.blogspot.com.au/2014/05/the-necessity-of-tobs.html
http://www.moyhu.blogspot.com.au/2014/07/tobs-pictured.html
http://www.moyhu.blogspot.com.au/2015/02/google-maps-and-ghcn-adjustments.html
http://www.moyhu.blogspot.com.au/2014/06/noaa-ghcn-station-portal.html
http://www.moyhu.blogspot.com.au/2015/02/homogenisation-makes-little-difference.html
http://www.moyhu.blogspot.com.au/2015/02/breakdown-of-effects-of-ghcn-adjustments.html

References

NOAA

1   Menne, Matthew J., and Claude N. Williams Jr. "Detection of undocumented changepoints using multiple test statistics and composite reference series." Journal of Climate 18.20 (2005): 4271-4286.

2   Menne, Matthew J., and Claude N. Williams Jr. "Homogenization of temperature series via pairwise comparisons." Journal of Climate 22.7 (2009): 1700-1717.

3   Menne, Matthew J., Claude N. Williams Jr, and Russell S. Vose. "The US Historical Climatology Network monthly temperature data, version 2." Bulletin of the American Meteorological Society 90.7 (2009): 993-1007.

4   Menne, Matthew J., Claude N. Williams, and Michael A. Palecki. "On the reliability of the US surface temperature record." Journal of Geophysical Research: Atmospheres (1984–2012) 115.D11 (2010).

5   Lawrimore, Jay H., et al. "An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3." Journal of Geophysical Research: Atmospheres (1984–2012) 116.D19 (2011).

6   Williams, Claude N., Matthew J. Menne, and Peter W. Thorne. "Benchmarking the performance of pairwise homogenization of surface temperatures in the United States." Journal of Geophysical Research: Atmospheres (1984–2012) 117.D5 (2012).

7   Quayle, Robert G., et al. "An operational near-real-time global temperature index." Geophysical research letters 26.3 (1999): 333-335.

8   Peterson, Thomas C., et al. "Homogeneity adjustments of in situ atmospheric climate data: a review." International Journal of Climatology 18.13 (1998): 1493-1517.

9   Smith, Thomas M., et al. "New surface temperature analyses for climate monitoring." Geophysical Research Letters 32.14 (2005).

10   Aguilar, Enric, et al. "Guidance on metadata and homogenization." WMO TD 1186 (2003): 53.

11   Peterson, Thomas C., et al. "Global Historical Climatology Network (GHCN) quality control of monthly temperature data." International Journal of Climatology 18.11 (1998): 1169-1179.

12   Homogenization code here.

GISS

13   Hansen, J., et al. "GISS analysis of surface temperature change." Journal of Geophysical Research: Atmospheres (1984–2012) 104.D24 (1999): 30997-31022.

14   Hansen, J., et al. "Current GISS Global Surface Temperature Analysis." NASA Goddard Institute for Space Studies, New York (2010).

15   GISTEMP code here.

Hadley/UEA

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

17   Jones, P.D., New, M., Parker, D.E., Martin, S. and Rigor, I.G., 1999: Surface air temperature and its variations over the last 150 years. Reviews of Geophysics 37, 173-199

18   Jones, P.D., Osborn, T.J., Briffa, K.R., Folland, C.K., Horton, B., Alexander, L.V., Parker, D.E. and Rayner, N.A., 2001: Adjusting for sampling density in grid-box land and ocean surface temperature time series. J. Geophys. Res. 106, 3371-3380

19   Jones, P.D., Lister, D.H., Osborn, T.J., Harpham, C., Salmon, M. and Morice, C.P., 2012: Hemispheric and large-scale land surface air temperature variations: an extensive revision and an update to 2010. Journal of Geophysical Research 117.

20   Kennedy J.J., Rayner, N.A., Smith, R.O., Saunby, M. and Parker, D.E., 2011: Reassessing biases and other uncertainties in sea-surface temperature observations measured in situ since 1850 part 2: biases and homogenisation. Journal of Geophysical Research 116.

21   Osborn, T.J. and Jones, P.D., 2014: The CRUTEM4 land-surface air temperature data set: construction, previous versions and dissemination via Google Earth. Earth System Science Data 6, 61-68.

22   Morice, C.P., Kennedy, J.J., Rayner, N.A. and Jones, P.D., 2012: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 dataset. Journal of Geophysical Research, 117

23   Rayner, N.A., P. Brohan, D.E. Parker, C.K. Folland, J.J. Kennedy, M. Vanicek, T. Ansell and S.F.B. Tett, 2006: Improved analyses of changes and uncertainties in marine temperature measured in situ since the mid-nineteenth century: the HadSST2 dataset. J. Climate, 19, 446-469