Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends.
UPDATE
Reconciling global temperature series

Kevin Cowtan, Robert G. Way

April 4th, 2014

1 Reconciling global temperature series

This update document provides a preliminary investigation into differences between the NASA GISTEMP global temperature reconstruction of Hansen *et al.* (2010) and the new HadCRUT4-based temperature reconstruction introduced in Cowtan and Way (2014) (henceforth CW14) and subsequent updates of 06/01/2014 and 14/02/2014. The CW14 infilled temperate reconstruction should be comparable to GISTEMP, since both provide essentially global coverage and both reconstruct air temperatures over sea ice in a similar manner. However the CW14 data show a trend of approximately 0.03°C/decade greater than GISTEMP over a period of 16 years.

As with previous documents, this work will focus primarily on temperature variations over the period 1997/01-2012/12. This period is of particular interest because it features rapid warming in the Arctic (Comiso and Hall, 2014), coupled with a cooling trend in the boreal mid latitudes (Cohen *et al.*, 2012). As established in CW14, rapid warming in the sparsely sampled Arctic provides a significant challenge for existing temperature averaging algorithms. However the same regional trends also present serious challenges for station homogenization algorithms, which depend on the assumption that climate trends are spatially correlated over moderate distances.

The study period therefore provides a pathological test for existing global mean surface temperature algorithms, and it is therefore unsurprising that different versions of the temperature record show divergence over this period. This divergence is however confusing to non-specialist users of temperature data, whether they be scientists with other specializations or members of the wider public audience.

Our initial speculation that the principal difference between GISTEMP and CW14 arose from the choice of sea surface temperature dataset was incorrect: the majority of the difference is in the Arctic, and arises from differences in the input land temperature data from meteorological stations. The GISTEMP record is based on the GHCN-Monthly Version 3 station data (Lawrimore *et al.*, 2011). There is some evidence that the homogenization adjustment algorithm used in the GHCN station data is attempting to eliminate some of the rapid Arctic warming over the study period.

2 Differences between the GISTEMP and CW14 reconstructions

An initial comparison between GISTEMP and the CW14 version 2 long kriging reconstruction was performed to determine whether the differences were geographically localized. Both gridded datasets were expanded to 1x1 degree grids and compared in the form of temperature trend maps. (The 1x1 degree grid is used by a number of other datasets and so is convenient for comparison). The difference in trend over the study period between GISTEMP and the CW14 is shown in Figure (U1). The most obvious differences are in the Arctic, primarily in the Barents and Kara seas north of Russia, and also in the Beaufort Sea north of Canada and Alaska. The region north of 60N accounts for about 2/3 of the difference in trend between GISTEMP and CW14 during the study period. A few rogue cells are also visible arising from CRUTEM4 stations which require homogenization adjustments due to changes in station equipment, location or operating practices.

A similar comparison was performed between the GISTEMP data and the recently released Berkeley land-ocean dataset (Rohde *et al.*, 2013). The results are shown in Figure (U2). The large differences in the Arctic are still apparent, however the lower latitude outliers are largely eliminated, providing support for the automated homogenization approaches used by Berkeley and GHCN. Some differences are apparent around Antarctica.

A comparison of Arctic temperatures between the GISTEMP, CW14, and Berkeley thermometer records, as well as the AVHRR satellite radiometer record Comiso and Hall (2014) and the MERRA reanalysis Rienecker *et al.* (2011) is shown in Figure (U3). All the datasets support higher Arctic temperatures than GISTEMP since 2006. AVHRR and MERRA both suggest that 2005 was warmer and 2011 cooler than the spatially interpolated thermometer records of CW14 and Berkeley. MERRA shows rapid warming over the study period but slower warming over the years prior to 1997.

What is the cause of these differences? The affected areas are covered by sea ice for a significant part of the year, and all three datasets reconstruct air temperatures over sea ice from the nearest land-based air temperature readings, suggesting that differences in the land temperature data themselves may be responsible for the differences.

GISTEMP, CW14 and Berkeley all depend on different weather station datasets:

- GISTEMP uses the GHCN-monthly dataset (Smith *et al.*, 2008), augmented by some additional stations. In the GHCN data the weather station data are automatically homogenized by pairwise comparison of nearby weather stations (Williams *et al.*, 2012). High latitude coverage is limited with only 3482 station months of data north of 70N in the study period.

- CW14 use the CRUTEM4 station data (Jones *et al.*, 2012) with an ad-

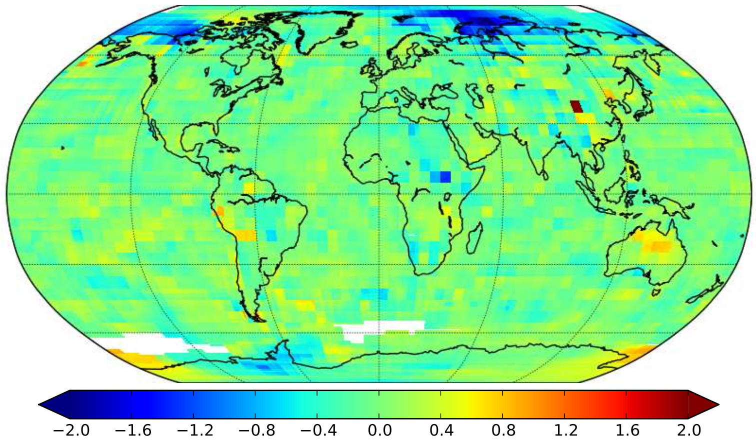

Figure U1: Temperature trend differences between GISTEMP and CW14 kriging reconstruction (v2) on the period 1997/01-2012/12 (i.e. GISTEMP minus CW14). Units are °C/decade.

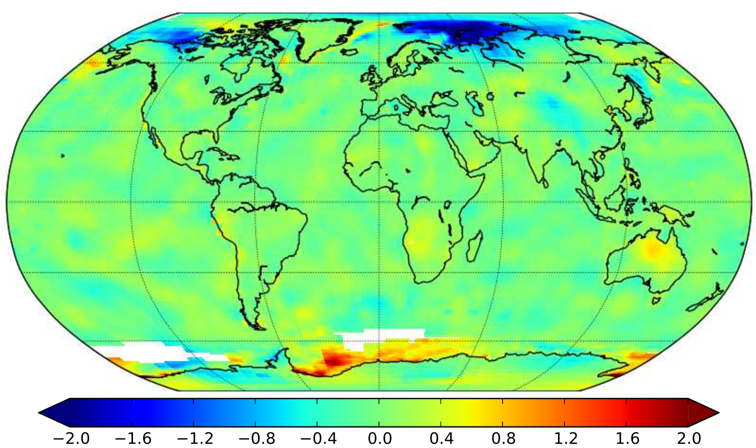

Figure U2: Temperature trend differences between GISTEMP and Berkeley land-ocean reconstruction on the period 1997/01-2012/12 (i.e. GISTEMP minus Berkeley). Units are °C/decade.

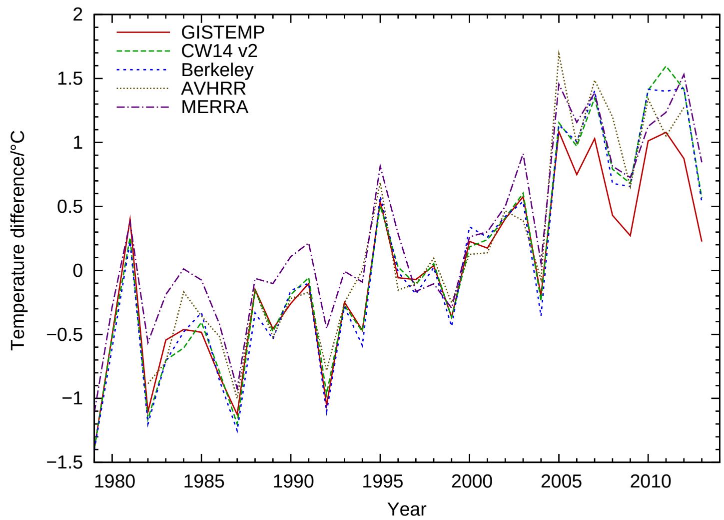

Figure U3: Temperature series for the region north of 64N from the GISTEMP, CW14 and Berkeley thermometer records, the AVHRR satellite record and the MERRA reanalysis. Records are aligned on the first 5 years of the study period (i.e.1997-2001) to enable the comparison of trends post-1997.

dition adjustment for the urban heat island effect applied as part of the HadCRUT4 ensemble calculation (Morice *et al.*, 2012). The data are not subject to automated homogenization, however some remediation has been carried out by the regional data providers or by the Climatic Research Unit. While the dataset includes fewer stations than GHCN in total, the CRUTEM4 update added many high latitude stations. As a result Arctic coverage is much better than GHCN with 13140 station months of data north of 70N in the study period, although there is some drop in coverage after 2008.

- Berkeley Earth use a more extensive list of stations than the other records, and their temperature averaging algorithm does not require that stations were active during the same baseline period. While some of the additional stations only provide short records, an automated homogenization procedure allows even fragmentary records to be aligned to a local climatology while detecting inhomogeneities in the individual records in the process.

Arctic trend differences between GISTEMP and five other temperature reconstructions are shown in Figure (U4). Comparisons are made with the CW14 v2 long kriging reconstruction, the CW14 hybrid with the UAH satellite data, the Berkeley Earth reconstruction, the MERRA reanalysis data and the AIRS satellite radiometer data (AIRS Science Team/Joao Texeira, 2013) (the latter covering only part of the period). The Barents/Kara and Beaufort differences are clearly visible in all of the comparisons. The MERRA reanalysis also shows a large region of faster warming on the Chukchi side of the central Arctic (showing up as a cool region in the difference map), counterbalanced by a region of slower warming off the coast of Greenland (which appears as warm in the difference map). Neither of these features, if real, would be captured by the thermometer records since there are no land-based weather stations in these regions. The CW14 hybrid with UAH offers weak support for the region of faster warming towards the Chukchi sea, however since MERRA assimilates the satellite data this is not an independent result.

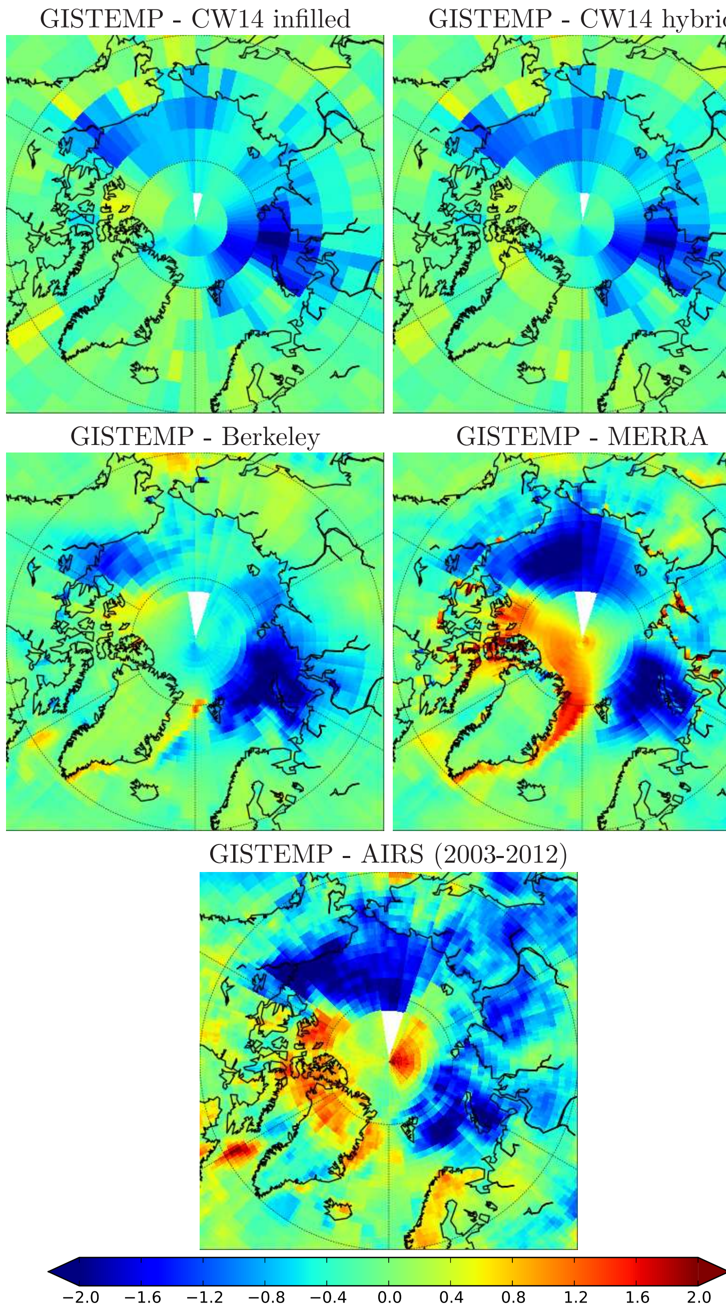

Figure U4: Temperature trend differences between GISTEMP and various temperature reconstructions on the period 1997/01-2012/12 (or 2003/01-2012/12 for AIRS). Units are °C/decade.

3 GHCN homogenization adjustments

The GHCN station homogenization algorithm (Williams *et al.*, 2012) uses multiple pairwise comparisons of nearby stations to detect discontinuities in station records, which may be due to changes in station location, equipment or operating practices. Simple adjustments, such as the application of a constant offset, are made to minimize such discontinuities. This approach has proven effective in identifying the impacts of station moves, instrument changes, time of observation and urbanization in the densely sampled 20th century US record (Williams *et al.*, 2012; Hausfather *et al.*, 2013). However recent Arctic warming presents two problems not present in the US case: firstly the station network in the high Arctic is sparse, and secondly the Arctic has been warming rapidly at the same time that the boreal mid latitudes have shown a cooling trend, especially over eastern Russia (Cohen *et al.*, 2012), illustrated in Figure (U5). The close proximity of regions of warming and cooling on both the Eurasian and Alaskan Arctic coasts mean that it is possible for neighbouring stations to show a very different temperature trends. Automated homogenization could potentially introduce unnecessary adjustments to reconcile these trends.

The impact of the homogenization calculation on the station data was estimated by feeding the GHCN data into a CRUTEM-like anomaly and gridding calculation. This calculation was performed for both the raw (i.e. pre-homogenization) and adjusted GHCN data. The difference between the resulting raw and adjusted map series was determined, and the trend in the difference over the 16 year study period plotted in Figure (U6). The results show a systematic pattern of downward temperature adjustments in Arctic stations over the study period, consistent with the hypothesis that the homogenization algorithm is trying to 'correct' for the rapid Arctic warming over that period. The adjustments are with a few exceptions (e.g. on the Chukchi peninsula) located in the regions where GISTEMP differs from the other temperature series.

The GHCN station data were reviewed to determine the stations responsible for the high-latitude adjustments in Figure (U6). Comparisons were made against the MERRA data for the nearest 1x1 grid cell to provide an indication of the plausibility of the adjustment. The adjustments were also compared with the adjustments from the Berkeley algorithm. The Berkeley data were also investigated to identify those neighbouring stations which most influence the homogenization process. A station-by-station summary is provided in Appendix A.

Two stations in the Kara sea and one in the Beaufort sea have homogenization adjustments which appear to be inconsistent with the MERRA data and with additional stations available in the Berkeley dataset. The affected stations are GMO IM. E. T., OSTROV VIZE, and BARTER ISLAND; labeled 3, 4 and 8 in Figure (U6) . The map comparisons suggest that further stations are affected, however these could not be identified in the preliminary station survey.

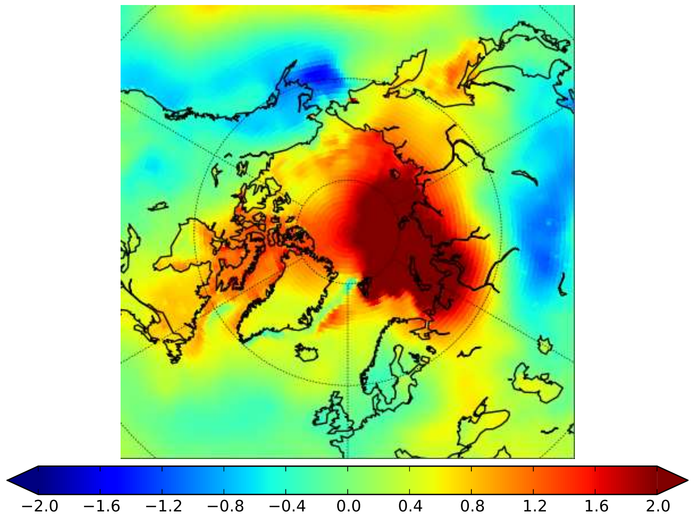

Figure U5: Trend on the period 1997/01 and 2012/12 in the Berkeley temperature data. The contrast between the Barents/Kara warming and the cooling trend over northern Eurasia is very pronounced. A weaker contrast is present on the Beaufort coast. Units are °C/decade.

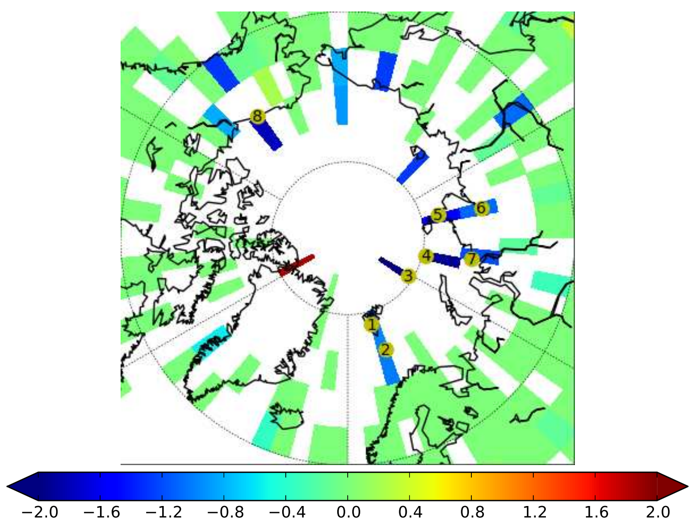

Figure U6: Difference in trend on the period 1997/01 and 2012/12 between the GHCN adjusted and raw data for cells with more than 60 months of data. Units are °C/decade. Numbers identify stations which are investigated in Appendix A, as follows: 1 - SVALBARD LUFTHAVEN, 2 - BJORNOYA, 3 - GMO IM. E.T., 4 - OSTROV VIZE, 5 - GMO IM.E.K F, 6 - HATANGA, 7 - OSTROV DIKSON, 8 - BARTER ISLAND

4 Detailed attribution of trend differences

The GISTEMP, CW14 and Berkeley calculation differ in a number of respects which could also contribute to the differing estimates of Arctic warming. In addition to using different weather station data, the 3 records also differ in station alignment algorithm, extrapolation algorithm, SST dataset and the method by which sea ice cells are identified. Any of these could have an impact on the Arctic temperature reconstruction.

Parts of the GISTEMP algorithm were therefore re-implemented to enable their effect on the temperature reconstruction to be evaluated. A simple temperature reconstruction algorithm was developed with the following steps:

- Station records were aligned using the common anomaly method (Jones *et al.*, 2012) using the period 1951-1980 as a baseline. (This is a deviation from the GISTEMP approach.)
- The temperature records were then gridded on an equal area grid with 2° sampling at the equator. Multiple station records in a single cell were averaged.
- Extrapolation was performed using a conical smoothing kernel with a radius of 1200km.
- The land-ocean reconstruction was created by blending the extrapolated land data with the ERSST ocean data (Smith *et al.*, 2008) according to the proportion of land and ocean in the cell. If the ERSST cell was missing (e.g. for sea ice), the land temperature was used.

To test the algorithm it was first applied to the GHCN adjusted data used in GISTEMP. The difference between GISTEMP and the resulting temperature reconstruction is very small in both the temperature series and the geographical distribution of trends - Figure (U7a). The agreement confirms that our implementation of the GISTEMP algorithm is adequate for attribution purposes.

Next the calculation was repeated substituting the raw data for the adjusted data for the 3 problem stations identified previously - Figure (U7b). The resulting difference only covers some of the region over which the temperature reconstructions disagree.

Next the calculation was repeated using the CRU station records, filtered to include only records in cells where a GHCN observation is available. For cells with a GHCN observation but no CRU observation, the nearest available CRU data were used. The difference between GISTEMP and this reconstruction is shown in Figure (U7c), and accounts for more of the difference between GISTEMP and CW14.

Finally the calculation was repeated using the full CRU temperature dataset. The resulting trend differences are shown in Figure (U7d), and are similar to the differences between the GISTEMP and CW14 reconstructions.

Trends for the period 1997/01-2012/12 for the region north of 64N are given in Table (U1). When using the GHCN adjusted data the trend agrees with

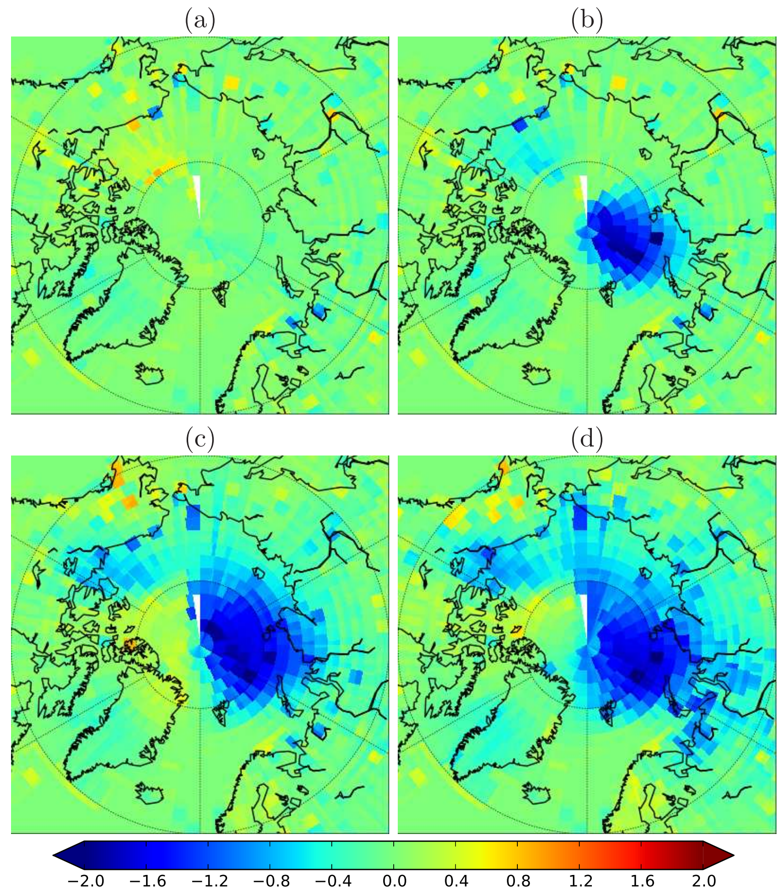

Figure U7: Temperature trend differences between GISTEMP and various temperature reconstructions on the period 1997/01-2012/12. Units are °C/decade. (a) Using the GHCN adjusted data. (b) Using the GHCN adjusted data with 3 stations reverted. (c) Using the CRU station data at GHCN sites only. (d) Using the full CRU data.

GISTEMP, and when using the CRU data the trend matches CW14. The intermediate reconstructions suggest that the 3 problem stations account for nearly 40% of the difference between GISTEMP and CW14. Other unidentified differences in the station data at locations present in both datasets also contribute nearly 40%. The remaining difference arises from stations present in CRU but absent in GHCN.

All of the difference between GISTEMP and CW14 in the Arctic can be accounted for by differences in the input station data. The station data are visualized in Figure (U8). GHCN and CRU both show plausible geographical variations in trend so neither can be immediately rejected as suffering from

Dataset	Trend (°C/decade)
GISTEMP	0.75
CW14	1.16
GHCN adjusted	0.74
... except 3 stations	0.90
CRU data, GHCN coverage	1.05
CRU data, full coverage	1.16

Table U1: Temperature trends for the region north of 64N in °C/decade for GISTEMP, CW14 and the four reconstructions from Figure (U7). Trends are on the period 1997/01-2012/12.

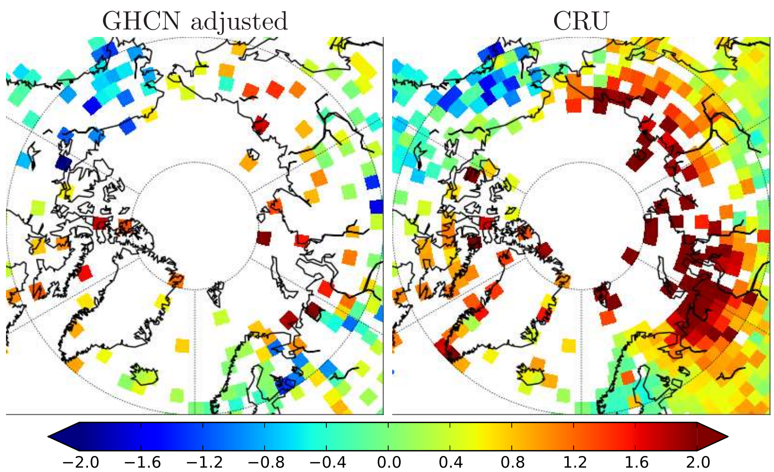

Figure U8: Temperature trends on the period 1997/01-2012/12 for the input station data. Units are °C/decade.

obvious homogenization problems. However temperature trends for the Arctic islands and some coastal regions differ substantially between the datasets, suggesting homogenization issues for one of the datasets.

The more extensive station inventory of the CRU data is clearly visible. The presence of consistent coordinated inhomogeneities across many geographically dispersed stations is unlikely, and so it seems more plausible that there are additional homogenization problems in the GHCN data not identified in the station survey. The agreement between CW14, Berkeley, satellite and reanalysis data also support this hypothesis.

If the differences between the GHCN adjusted data and the other station datasets are indeed due to inappropriate homogenization adjustments, then it is likely that the problem will be resolved with the introduction of GHCN-M version 4, which will include a more extensive station inventory from the International Surface Temperature Initiative "Stage 3" product available from http://www.surfacetemperatures.org (Thorne *et al.*, 2011). (The Berkeley records for the most challenging stations are generally supported by neighbouring stations not present in GHCN v3, but which are likely to be included in GHCN v4.)

In summary, re-implementation of the GISTEMP algorithm shows that the difference between GISTEMP and CW14 Arctic temperature trends can be almost completely explained by differences in the input station data. The choice of algorithm and SST dataset is essentially irrelevant. When provided with the same station data the GISTEMP algorithm produces similar Arctic temperatures over the study period to the algorithm of CW14, and the results are also similar to those of Berkeley Earth.

5 Discussion

The starting point of our investigations into both coverage and homogenization bias was the question 'why do the datasets differ?'. Initial investigations lead to the conclusion that coverage was the most important issue, and on this basis we expected that addressing the coverage bias in HadCRUT4 would bring it into agreement with GISTEMP. Instead we were surprised to discover that our data showed more rapid recent warming than GISTEMP. Investigating the cause of the differences in turn lead to the conclusion that the differences arise primarily from the input weather station data rather than differences in methodology or SST data. The GHCN homogenizations of Arctic stations are a likely source of much of the divergence.

These results support the conclusion of CW14 that the global temperature trend over the study period lies outside the range bounded by the existing datasets. While the remaining differences could have been dismissed as structural uncertainty, understanding those differences led to the elucidation of potential biases. As a result we suggest that it is important to attempt to understand where and why the different temperature series differ rather than dismissing the differences as uncertainty.

The study period represents a severe test of algorithms for global mean temperature estimation due to both the rapid localised climate change in a region of sparse observations, and also the high spatial gradients between regions of warming and cooling. Given that this behaviour appears to be unprecedented since the 1950s when weather station coverage first approached current levels, it is unsurprising that problems should arise. However for the same reasons the study period provides an important test for the development and evaluation of more robust methods. A number of new reanalysis and radiometer datasets provide additional tools for validating the results. Manual reconstructions of key high latitude stations, such as those of Bromwich *et al.* (2013) and Nordli *et al.* (2014) may also be important.

Finally we note that the new Berkeley land-ocean dataset provides good coverage and the homogenization results appear to be robust: obvious low latitude station homogenization problems are remedied without attenuating the recent Arctic warming. On this basis we consider it to be an important addition to the understanding of the historical temperature record and we recommend its use alongside the more established datasets for the analysis of historical temperature changes.

Acknowledgements

We would like to thank Matthew Menne and Claude Williams for helpful discussions relating to this problem. We are also grateful to Zeke Hausfather for comments and Steve Mosher for help with the AIRS data.

References

AIRS Science Team/Joao Texeira. 2013. Aqua airs level 3 monthly standard physical retrieval (airs+amsu),version 006, greenbelt, md, usa:nasa goddard earth science data and information services center (ges disc), accessed at doi:10.5067/aqua/airs/data319.

Bromwich DH, Nicolas JP, Monaghan AJ, Lazzara MA, Keller LM, Weidner GA, Wilson AB. 2013. Central west antarctica among the most rapidly warming regions on earth. *Nature Geoscience* **6**(2): 139–145.

Cohen JL, Furtado JC, Barlow M, Alexeev VA, Cherry JE. 2012. Asymmetric seasonal temperature trends. *Geophysical Research Letters* **39**(4).

Comiso JC, Hall DK. 2014. Climate trends in the arctic as observed from space. *Wiley Interdisciplinary Reviews: Climate Change* .

Cowtan K, Way RG. 2014. Coverage bias in the hadcrut4 temperature series and its impact on recent temperature trends. *Quarterly Journal of the Royal Meteorological Society* .

Hansen J, Ruedy R, Sato M, Lo K. 2010. Global surface temperature change. *Reviews of Geophysics* **48**(4): RG4004.

Hausfather Z, Menne MJ, Williams CN, Masters T, Broberg R, Jones D. 2013. Quantifying the effect of urbanization on us historical climatology network temperature records. *J. Geophys. Res.: Atmospheres* doi:10.1029/2012JD018509. In press.

Jones P, Lister D, Osborn T, Harpham C, Salmon M, Morice C. 2012. Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010. *J. Geophys. Res.* **117**(D5).

Lawrimore JH, Menne MJ, Gleason BE, Williams CN, Wuertz DB, Vose RS, Rennie J. 2011. An overview of the global historical climatology network monthly mean temperature data set, version 3. *Journal of Geophysical Research* **116**(D19): D19 121.

Lindsay R, Wensnahan M, Schweiger A, Zhang J. 2014. Evaluation of seven different atmospheric reanalysis products in the arctic. *Journal of Climate* **in press**(2014).

Morice CP, Kennedy JJ, Rayner NA, Jones PD. 2012. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The hadcrut4 data set. *J. Geophys. Res.* **117**(D8): D08 101.

Nordli Ø, Przybylak R, Ogilvie AE, Isaksen K. 2014. Long-term temperature trends and variability on spitsbergen: the extended svalbard airport temperature series, 1898-2012. *Polar Research* **33**.

Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim GK, Bloom S, Chen J, Collins D, Conaty A, da Silva A, Gu W, Joiner J, Koster RD, Lucchesi R, Molod A, Owens T, Pawson S, Pegion P, Redder CR, Reichle R, Robertson FR, Ruddick AG, Sienkiewicz M, Woollen J. 2011. Merra: Nasa's modern-era retrospective analysis for research and applications. *Journal of Climate* **24**(14): 3624–3648.

Rohde R, Muller RA, Jacobsen R, Perlmutter S, Rosenfeld A, Wurtele J, Curry J, Wickham C, S M. 2013. Berkeley earth temperature averaging process. *Geoinformatics & Geostatistics: An Overview* doi:10.4172/gigs.1000103.

Smith TM, Reynolds RW, Peterson TC, Lawrimore J. 2008. Improvements to noaa's historical merged land-ocean surface temperature analysis (1880-2006). *J. Climate* **21**(10): 2283–2296.

Thorne PW, Willett KM, Allan RJ, Bojinski S, Christy JR, Fox N, Gilbert S, Jolliffe I, Kennedy JJ, Kent E, *et al.* 2011. Guiding the creation of a comprehensive surface temperature resource for twenty-first-century climate science. *Bulletin of the American Meteorological Society* **92**(11): ES40–ES47.

Williams CN, Menne MJ, Thorne PW. 2012. Benchmarking the performance of pairwise homogenization of surface temperatures in the united states. *J. Geophys. Res.* **117**(D5): D05 116.

A High latitude station homogenizations

While the impact of homogenization on the GHCN-derived time series and the geographical distribution of homogenizations over the study period both suggest that the GHCN homogenization algorithm may be trying to suppress the rapid high latitude warming signal, it is also possible that there may be real homogenization issues with inaccessible high latitude stations. To assess this possibility, the GHCN homogenizations were compared with results from a similar homogenization algorithm implemented by the Berkeley Earth Surface Temperature project, and also with the MERRA reanalysis dataset (Rienecker *et al.*, 2011). Since MERRA does not assimilate weather station temperature observations it provides an independent estimate of land surface air temperature. Lindsay *et al.* (2014) report that it outperforms other reanalysis products in the Arctic.

When comparing with MERRA, a time series is extracted for the MERRA grid cell containing the station. The difference between both the GHCN adjusted and unadjusted series and the MERRA series are plotted. If only one of these shows a discontinuity, then this provides support for the other. For this preliminary analysis no formal breakpoint analysis or significance testing have been performed.

The comparisons in this section are for the stations which are most strongly implicated in the divergence between GISTEMP and CW14. GHCN adjustments to other stations, such as Alert N.W.T. and stations on the Chukchi peninsula appear to be well founded and have been omitted. The most problematic adjustments are for the stations GMO IM. E.T. (POLARGMO) and OSTROV VIZE in the Barents/Kara region, and BARTER ISLAND in the Beaufort region - Berkeley and CRU have additional stations in these areas. The stations around Svalbard and Severnaya Zemlya are on the borders of region of disagreement and the correct adjustments for these stations are less clear.

The GHCN data used in this report is from GHCN-M v3.2.2.20140125. More recent versions may differ due to additional data or updated algorithms.

A.1 SVALBARD LUFTHAVEN (GHCN: 63401008000)

GHCN record:

ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/products/stnplots/6/63401008000.gif

Berkeley record:

http://berkeleyearth.lbl.gov/stations/159256

Nearby Berkeley stations:

- BARENCBURG http://berkeleyearth.lbl.gov/stations/94190
- SVEAGRUVA http://berkeleyearth.lbl.gov/stations/19168
- NY-ALESUND I http://berkeleyearth.lbl.gov/stations/157317
- NY-ALESUND II http://berkeleyearth.lbl.gov/stations/159257
- HORNSUND http://berkeleyearth.lbl.gov/stations/159254

GHCN apply a ∼1C downward adjustment to this station starting in March 2004. Berkeley do not apply a correction to the station record in this timeframe. The Berkeley record is supported by numerous nearby stations. MERRA does not support a break in early 2004, but does support one around 2006. MERRA may also support a possible discontinuity in early 2012 requiring an upward adjustment. Berkeley notes a station move in August 2012.

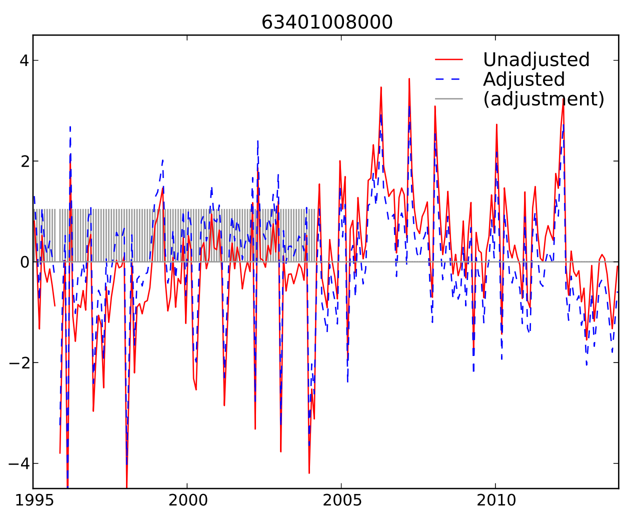

Figure U9: Difference in monthly temperatures between GHCN raw and adjusted data and MERRA.

A.2 BJORNOYA (GHCN: 63401028000)

GHCN record:

ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/products/stnplots/6/63401028000.gif

Berkeley record:

http://berkeleyearth.lbl.gov/stations/157315

Nearby Berkeley stations:

- HOPEN http://berkeleyearth.lbl.gov/stations/159253
- HORNSUND http://berkeleyearth.lbl.gov/stations/159254

GHCN apply a ∼1C downward adjustment to this station starting in March 2004. Berkeley apply a correction to the station record in September 2003. The Berkeley correction is supported by the HOPEN and HORNSUND stations 200-300km away. However MERRA strongly rejects this break. The HOPEN record contains a break in August 1998 - it is possible that BJORNOYA could be reconciled with both HOPEN and MERRA by shifting the BYORNOYA break back to 1998 and the HOPEN break to 2004.

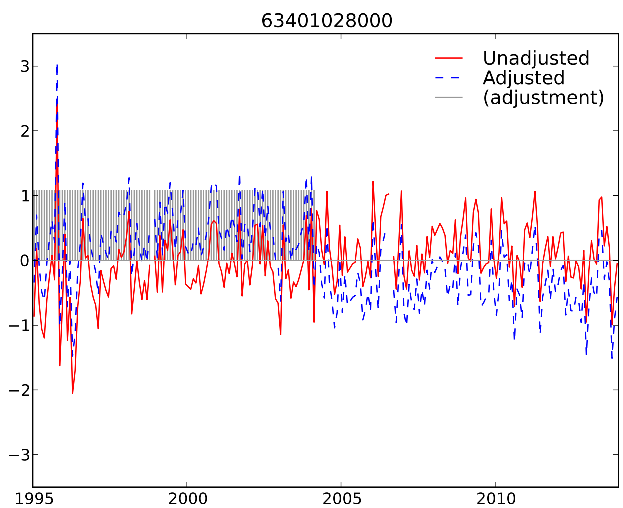

Figure U10: Difference in monthly temperatures between GHCN raw and adjusted data and MERRA.

A.3 GMO IM. E.T. (GHCN: 22220046000)

GHCN record:
ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/products/stnplots/2/22220046000.gif

Berkeley record:
http://berkeleyearth.lbl.gov/stations/169963

Nearby Berkeley stations:

- OSTROV VIZE http://berkeleyearth.lbl.gov/stations/169961

This station has exaggerated importance in the CW14 Arctic temperature reconstruction due to its location above 80N, which places it in a different row of cells to the other Eurasian Arctic stations, however this is counteracted somewhat by the limited number of months available. This is less of an issue in the GISTEMP and Berkeley records because extrapolation is based on the station position rather than the grid cell centre.

The data for this station is missing over the period 2000-2010. GHCN apply a large adjustment of ∼3C between observations before and after this period. The nearby station at OSTROV VIZE also receives a large adjustment in the GHCN data. Berkeley shows a much smaller adjustment, supported by the Berkeley record for OSTROV VIZE which features no adjustment until a station move in August 2011. MERRA rejects the large GHCN adjustment.

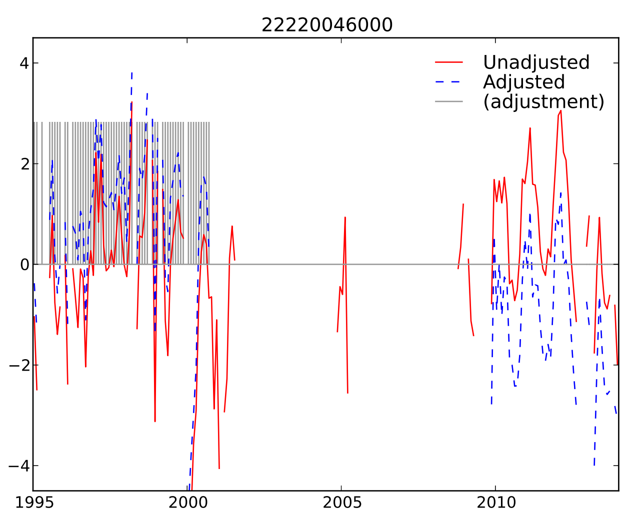

Figure U11: Difference in monthly temperatures between GHCN raw and adjusted data and MERRA.

A.4 OSTROV VIZE (GHCN: 22220069000)

GHCN record:
  ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/products/stnplots/2/22220069000.gif

Berkeley record:
  http://berkeleyearth.lbl.gov/stations/169961

Nearby Berkeley stations:

- GMO IM. E.T. http://berkeleyearth.lbl.gov/stations/169963
- MYS GOLOMIANNY / OSTROV P http://berkeleyearth.lbl.gov/stations/169962

GHCN shows a downward adjustment to this station of about 1C in Dec 2004, and a larger downward adjustment of 2C in Sep 2011 along with the rejection of observations from Oct 2009 to Sep 2011, presumably to maintain consistency with an adjustment in the previous station (GHCN:22220046000).

Berkeley does not support an adjustment in 2004 or 2009, although it does show an adjustment of nearly 2C in August 2011. As with the previous station MERRA rejects the adjustments.

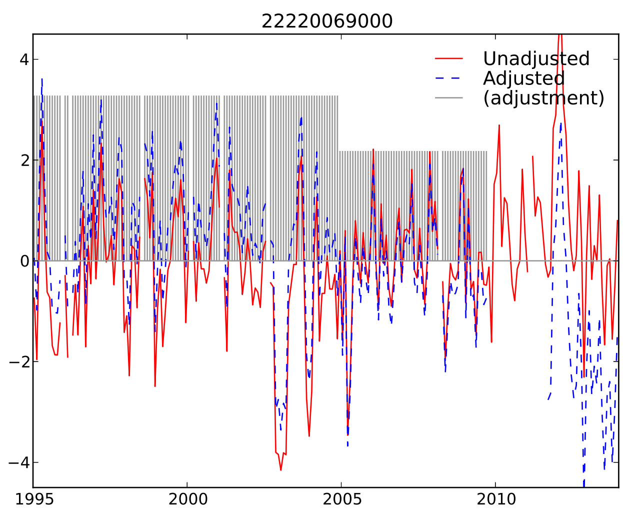

Figure U12: Difference in monthly temperatures between GHCN raw and adjusted data and MERRA.

A.5 GMO IM.E.K F (GHCN: 22220292000)

GHCN record:
ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/products/stnplots/2/22220292000.gif

Berkeley record:
http://berkeleyearth.lbl.gov/stations/169960

Nearby Berkeley stations:

- MYS GOLOMIANNY / OSTROV P http://berkeleyearth.lbl.gov/stations/169962
- CAPE STERLEGOVA http://berkeleyearth.lbl.gov/stations/175624
- OSTROV VIZE http://berkeleyearth.lbl.gov/stations/169961
- KHATANGA http://berkeleyearth.lbl.gov/stations/169947

GHCN shows a downward adjustment to this station in Dec 2005, and a further adjustment of over 1C in Oct 2009. Berkeley shows similar adjustment in 2005, but places the second adjustment in Oct 2011 associated with a station move. MERRA supports the earlier date for the second adjustment and a rather larger downward adjustment for the 2005-2009 period.

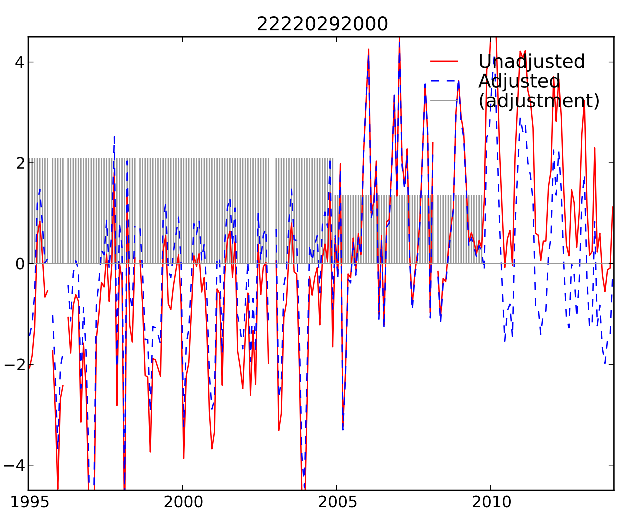

Figure U13: Difference in monthly temperatures between GHCN raw and adjusted data and MERRA.

A.6 HATANGA (GHCN: 22220891000)

GHCN record:
ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/products/stnplots/2/22220891000.gif

Berkeley record:
http://berkeleyearth.lbl.gov/stations/169947

Nearby Berkeley stations:

- VOLOCHANKA,AMSG http://berkeleyearth.lbl.gov/stations/169943
- DZALINDA-1 http://berkeleyearth.lbl.gov/stations/169937
- OLENEK http://berkeleyearth.lbl.gov/stations/169916
- CAPE STERLEGOVA http://berkeleyearth.lbl.gov/stations/175624

GHCN apply a downward adjustment of ∼1.5C to this station from January 2009. Berkeley show a much smaller adjustment in December 2005 supported by several nearby stations. MERRA tends to reject the large GHCN adjustment.

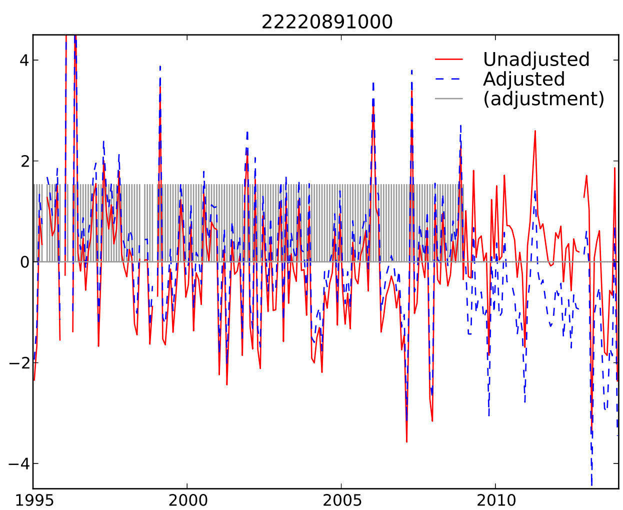

Figure U14: Difference in monthly temperatures between GHCN raw and adjusted data and MERRA.

A.7 OSTROV DIKSON (GHCN: 22220674000)

GHCN record:
ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/products/stnplots/2/22220674000.gif

Berkeley record:
http://berkeleyearth.lbl.gov/stations/169952

Nearby Berkeley stations:

- IM. M.V. POPOVA http://berkeleyearth.lbl.gov/stations/175623
- CAPE STERLEGOVA http://berkeleyearth.lbl.gov/stations/175624

GHCN apply a downward correction of ~1.5C from April 2008. Berkeley apply only a small adjustment in early 2005, and has support from nearby stations. MERRA supports the GHCN adjustment.

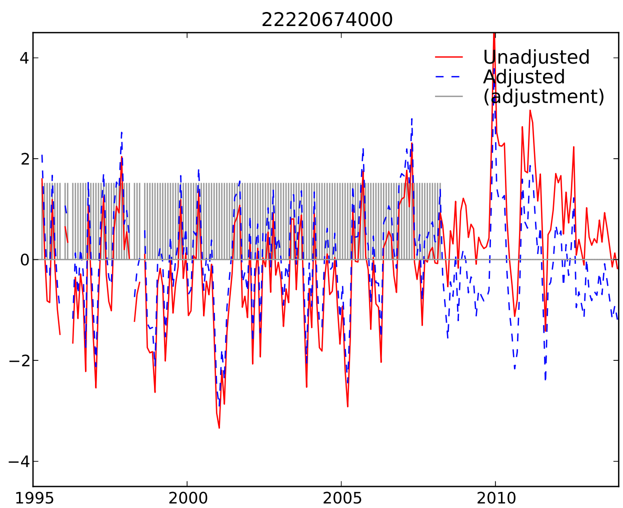

Figure U15: Difference in monthly temperatures between GHCN raw and adjusted data and MERRA.

A.8 BARTER ISLAND (GHCN: 42570086000)

GHCN record:

ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/products/stnplots/4/42570086000.gif

Berkeley record:

http://berkeleyearth.lbl.gov/stations/167758

Nearby Berkeley stations:

- IVVAVIK PARK AUTO8 http://berkeleyearth.lbl.gov/stations/11097
- HERSCHEL ISLAND, YT http://berkeleyearth.lbl.gov/stations/153857
- SHINGLE POINT A, YT http://berkeleyearth.lbl.gov/stations/153846

GHCN shows a downward adjustment to this station of about 1.5C in Jan 2006. Berkeley shows only a very small adjustment in 2010. The Berkeley data are more complete and is supported by a significant number of local stations. MERRA is of limited use given the shortness of the record fragments, although the fragment from 2000 appears to be inconsistent with the adjustment.

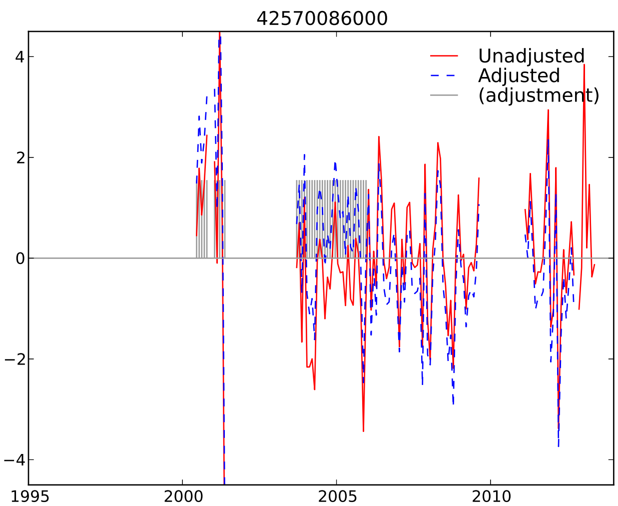

Figure U16: Difference in monthly temperatures between GHCN raw and adjusted data and MERRA.

B Impact of GHCN adjustments on global temperatures

The raw and adjusted GHCN temperature data were gridded and used as the basis for global temperature reconstructions using the method of CW14. The difference in the resulting global temperature series is shown in Figure (U17). The impact of the GHCN adjustments on a global temperature reconstruction is to lower temperatures by up to 0.02C over the past 5 years. The adjustments also increase temperatures by a smaller amount in the first few years of the 21st century.

This effect arises when the station data are used as the basis of a global temperature reconstruction (as in GISTEMP, Berkeley and CW14), because stations in sparse regions are upweighted to cover the missing regions. The effect on a non-global reconstruction such as the NOAA or HadCRUT temperature series is minimal. Nonetheless, claims that GHCN adjustments contribute to the warming trend over the satellite era are unfounded.

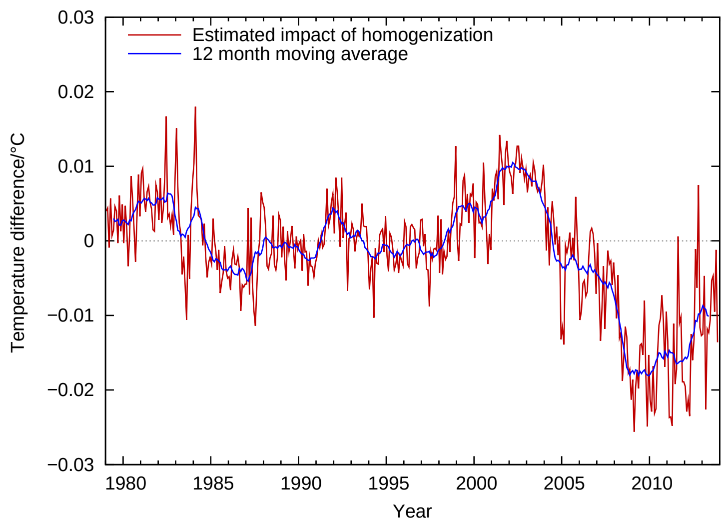

Figure U17: Difference in monthly global mean surface temperature estimate between global reconstructions from the GHCN adjusted and raw data.

C Seasonal trends in the Arctic

It is possible that differences in the allocation of either air or SST temperatures to Arctic cells could explain part of the divergence between the datasets. Since the biggest changes in sea ice cover have occurred in summer and autumn, this effect should appear as a different pattern of seasonal Arctic trends between the datasets. The seasonal trends for each of the temperature series are shown in Table (U2). The GISTEMP trends are consistently lower in all seasons, suggesting that changes in sea ice do not play a major role in the differences between the reconstructions.

	GISTEMP	CW14 kriging	Berkeley	MERRA
DJF	1.36	1.86	1.86	1.83
MAM	0.55	0.87	0.72	1.00
JJA	0.19	0.67	0.63	0.54
SON	0.87	1.21	1.31	1.11

Table U2: Seasonal temperature trends for the region north of 64N in °C/decade for GISTEMP and three other temperature reconstructions. Trends are on the period 1997/01-2012/12.