Influence of Urban Heating on the Global Temperature Land Average Using Rural Sites Identified from MODIS Classifications

Charlotte Wickham[1], Robert Rohde[2], Richard Muller[3,4], Jonathan Wurtele[3,4], Judith Curry[5], Don Groom[3], Robert Jacobsen[3,4], Saul Perlmutter[3,4], Arthur Rosenfeld[3].

[1] University of California, Berkeley, CA, 94720; currently at Dept. Statistics, Oregon St. University; [2] Novim Group, 211 Rametto Road, Santa Barbara, CA, 93104; [3] Lawrence Berkeley Laboratory, Berkeley, CA, 94720; [4] Department of Physics, University of California, Berkeley CA 94720, [5]Georgia Institute of Technology, Atlanta, GA 30332;. Correspondence for all authors should be sent to The Berkeley Earth Project, 2831 Garber Street, Berkeley CA, 94705.

Abstract

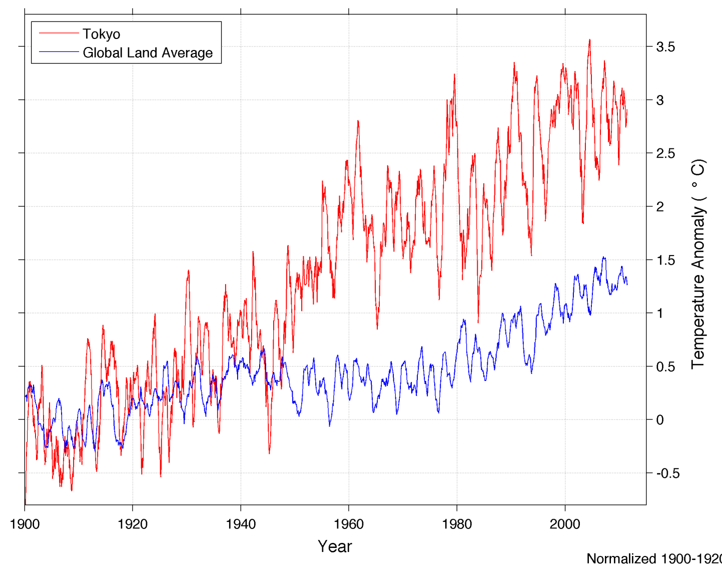
The effect of urban heating on estimates of global average land surface temperature is studied by applying an urban-rural classification based on MODIS satellite data to the Berkeley Earth temperature dataset compilation of 36, 869 sites from 15 different publicly available sources. We compare the distribution of linear temperature trends for these sites to the distribution for a rural subset of 15, 594 sites chosen to be distant from all MODIS-identified urban areas. While the trend distributions are broad, with one-third of the stations in the US and worldwide having a negative trend, both distributions show significant warming. Time series of the Earth's average land temperature are estimated using the Berkeley Earth methodology applied to the full dataset and the rural subset; the difference of these is consistent with no urban heating effect over the period 1950 to 2010, with a slope of -0.10 +- 0.24 / 100yr (95% confidence).

1. Introduction

The Urban Heat Island (UHI) effect describes the observation that temperatures in a city are often higher than in its rural surroundings. London was the first urban heat island to be documented (Howard ,1833) but since then many cities have been identified as urban heat islands (see Chandler,1976; Oke, 1974, 1979 and Arnfield , 2003). A well-known example is Tokyo where the temperature has risen much more rapidly in the city than in nearby rural areas: Fujibe (2011) estimates excess warming of almost $2^{o}C/100yr$ compared to the rest of Japan. The warming of Tokyo is dramatic when compared to a global average as seen in Fig.1. The UHI effect can be attributed to many physical differences between urban and rural areas, including absorption of sunlight, increased heat storage of artificial surfaces, obstruction of re-radiation by buildings, absence of plant transpiration, differences in air circulation, and other phenomena (Oke, 1982).

Urban areas are heavily overrepresented in the siting of temperature stations: less than 1% of the globe is urban but 27% of the Global Historical Climatology Network Monthly (GHCN-M) stations are located in cities with a population greater than 50,000. If the typical urban station exhibited urban heating of the magnitude of Tokyo this could result in a severe warming bias in global averages using urban stations. To avoid this bias the urban heating contribution to global temperature change should be isolated to the greatest extent possible.

Figure 1 Annual running mean of monthly temperatures at Tokyo compared to a global land average for 1900-2010



The goal of this paper is to evaluate the urban heat island contribution to the Berkeley Earth Surface Temperature global land average. Detailed analyses of average land temperature time series of the Earth's surface (T_{avg}) have been reported by three major teams: the NASA Goddard Institute for Space Science (GISS), the National Oceanographic and Atmospheric Administration (NOAA), and the collaboration between the Hadley Centre of the UK Met Office and the Climatic Research Unit of the University of East Anglia (HadCRU). They differ in the methods used to account for the effect of urban heating on their global averages. The conclusion of the three groups is that the urban heat island contribution to their global averages is much smaller than the observed global warming. The topic is not without controversy. We ask whether the presence of urban stations results in overestimates of warming in the Berkeley Earth Surface Temperature global land average.

The approach of the GISS team is to identify urban, “peri-urban” (near urban) and rural stations using satellite images of nighttime lights (Hansen et al., 2010). Urban and peri-urban stations are then adjusted by subtracting a two-part linear trend based on comparison to an average of nearby rural stations. The result of the adjustment on their global average is a reduction of about 0.01°C in warming over the period 1900 - 2009.

The NOAA group does not perform a specific urban adjustment in their most recent analysis, GHCN-M version 3. They use an automated pairwise comparison procedure (Menne & Williams, 2009) to make adjustments for documented and undocumented changes in station records, and expect that this process will remove most urban warming. When applied to the United States Historical Climatology Network, Menne et al. (2009) report that the average minimum temperature of the 30% most urban stations (based on population metadata) rises 0.06°C per century more than the more rural locations between 1895 – 2007.

The HadCRU group does not specifically model or adjust for urban warming (although some sites suspected to be influenced by urbanization are excluded from their analysis (Jones et al. 1986a, 1986b)). Instead, they include an estimate for the UHI effect when they give their uncertainty statement. In a recent analysis, (‘HadCRUT3’, Brohan et al. 2006) they add a one-sided one sigma uncertainty starting in 1900 and increasing linearly by 0.055°C per century. This value is based on a previous analysis of urban heating by Jones (1990).

Our approach most closely aligns with the studies of Jones et al. (1990), Peterson et al. (1999) and Parker (2004), in that two averages are constructed, one based on stations that should be free of urban heating and one using all station data. The difference between the two averages is examined for evidence that using all stations, including those suspected of containing urban heating, overestimates warming. Despite using different methods to identify a rural global average, all three concluded that the magnitude of the effect of urban heating on the global averages examined was small.

Other studies have examined the urban warming at a station level. Karl (1988) paired rural stations with nearby urban stations in the USHCN and found a warming bias that increases with increasing population of the city associated with the urban station. However, due to the small number of stations affected he concluded that the magnitude of the effect would be small in a US average. Peterson (2003) also compared urban and rural stations in the USA and found after careful consideration of inhomogeneities, station location, time of observation bias and instrumentation differences the apparent warm bias of urban stations was insignificant.

De Laat & Maurellis (2006) used industrial CO_2 emissions to classify stations into industrialized and non-industrialized stations. They found a significant increase in average temperature trend for industrialized stations. McKitrick & Micheals (2004, 2007) found significant correlations between the local trend in gridded averages and a number of social and economic indicators, and estimated one half of the observed global warming trend (over

1980 - 2002) might be due to these factors. Schmidt (2009) dismissed these studies as finding spurious correlations due to inadequate modeling of spatial correlation and that the robustness to alternative sources of data needed to be assessed. McKitrick & Nierenberg (2010) countered with an analysis designed to answer Schmidt's criticisms and confirmed their earlier findings.

The apparent contradiction of these studies is partly due to different areas of focus. De Laat & Maurellis (2006), McKitrick & Micheals (2004, 2007) and McKitrick & Nierenberg (2010) use the CRUTemp gridded product, which as mentioned, contains little effort to remove urban heating effects. McKitrick & Micheals (2004, 2007) and McKitrick & Nierenberg (2010) also focus on finding the heating signal in local trends rather than evaluating the effect on a global average.

We consider two sets of stations, a complete set and a set restricted to sites that are far from urban regions. To accomplish this we use the MODIS urban classification map (Schneider et al. 2009, 2010; described below) combined with our large collection of temperature stations. This is a larger set of stations than previous analyses have included. We first describe the datasets, and place the problem of estimating urban heating in context by conducting an investigation of the linear trends in this large set of temperature stations. Our primary analysis of the significance of site selection restricted to non-urban stations is then performed with the Berkeley Earth Temperature averaging procedure.

2. Data

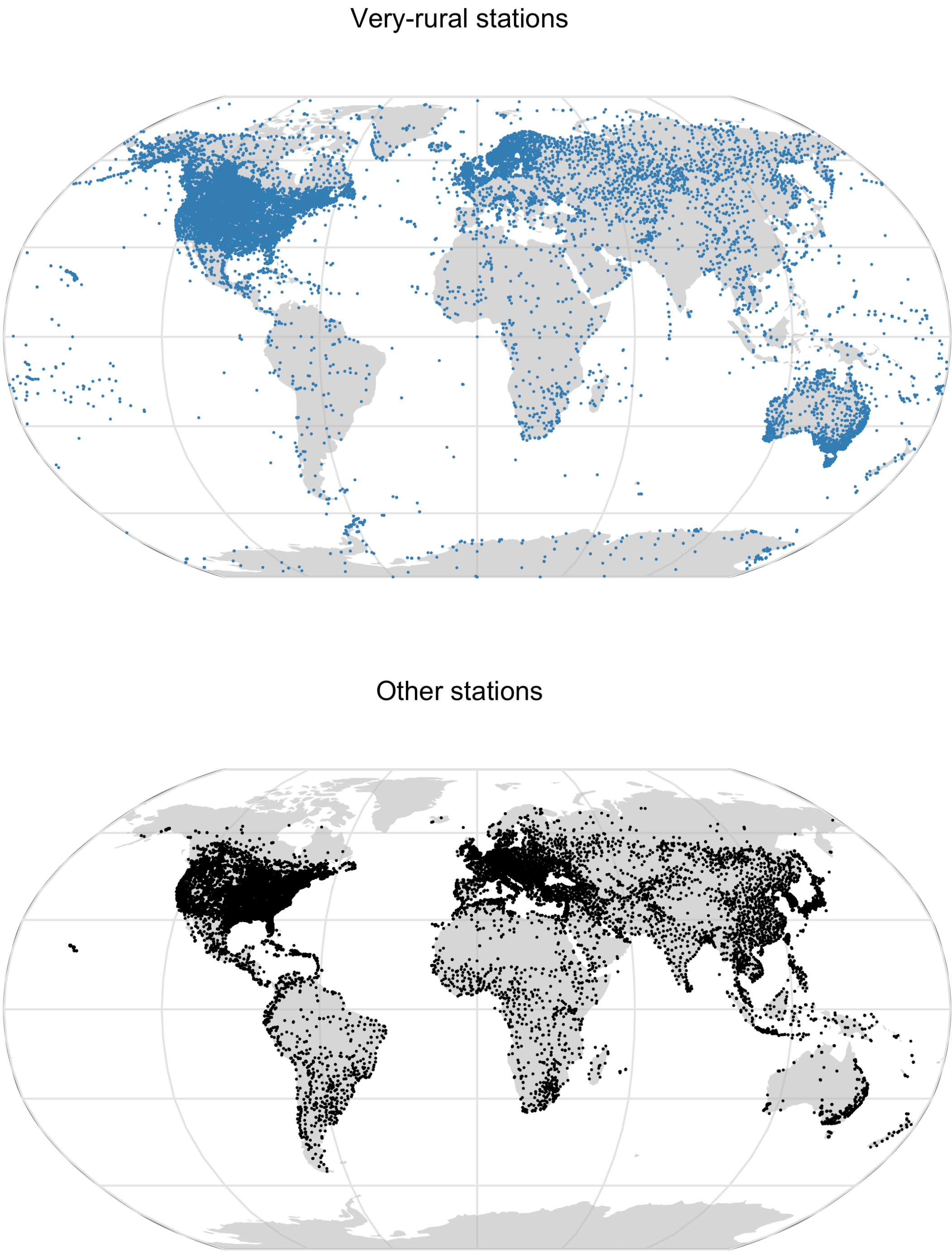
The analysis presented here is based on merged monthly average temperatures from the Berkeley Earth Surface Temperature Study dataset. This dataset consists of measurements from 36, 869 unique stations, which are merged from 15 preexisting data archives (the dataset and a description of the merging and filtering can be found at, http://berkeleyearth.org/dataset/). We classify these stations as rural or non-rural by comparing their locations with the MODIS 500m Global Urban Extent classification map (MOD500) of Schneider et al. (2009, 2010). Schneider et al. used Collection 5 MODIS 500-m resolution satellite imagery to classify land use as urban using supervised decision trees, a statistical learning algorithm that they trained using a set of sites with known land cover type. They define urban areas to be “places that are dominated by the built environment”. Urban heat islands are primarily a result of replacing the natural (soil, vegetation, etc.) surface of the land with buildings and artificial ground surfaces, which makes the MOD500 dataset potentially quite helpful in identifying built-up regions that may be subject to urban heating. It may provide a criterion that is less socio-economically biased than night lights data, therefore it offers an alternative to the approach used by GISS. The MOD500 map is available as a raster image, providing a binary classification (*urban* or *not urban*) for a global grid with pixels of size 15 arc-seconds. According to Potere et al. (2009) the MOD500 map outperforms other global urban maps in terms of predicting city size and per pixel agreement on a sample of known cities with population greater than 100,000.

Unfortunately, a portion of station locations in the Berkeley Earth merged dataset are reported only to the nearest tenth of a degree in latitude and longitude. This makes it impossible to identify each station as definitively urban or rural using the fine resolution MOD500 map. This imprecision in site location could yield a site which is urban being labeled as rural. An alternative, which we adopt here, is to analyze the urban-rural split in a different way. Rather than compare urban sites to non-urban, thereby explicitly estimating UHI effects, we split sites into very-rural and not very-rural. We defined a site as “very-rural” if the MOD500 map showed no urban regions within one tenth of a degree in latitude or longitude of the site. We expect these very-rural sites to be reasonably free from urban heating effects. Of the 36,869 sites, 15,594 were classified by this method as very-rural. The station locations and their classifications are displayed in Figure 2. Although the continental USA looks saturated with very-rural sites this is due to the density of stations in the USA and overplotting of points. In actuality 18% of the stations in the USA are classified as very-rural by our method.

We note that the imprecision in station locations also affects the GISS night lights analysis, with approximately 1/8th of the stations in their study also being positioned to only the nearest tenth of a degree. The GISS analysis (Hansen et al., 2010) does not explicitly address the possibility that station types might be misclassified due to geolocation uncertainties that far exceed to the 30 arcsecond resolution of the night lights maps.

The MOD500 map identifies urban areas circa 2001. We make the assumption that areas that are *not urban* in the MOD500 map, have always been *not urban*. This means a station classified as very-rural, is assumed to have been rural for the length of its station history. A corresponding continuity assumption would be inappropriate for areas that are *urban* in the MOD500 map, and this is one of the reasons we take the approach of identifying the stations that should be free of urban heating, rather than trying to identify stations that are subject to urban heating. A strict definition of rural allows us to build a global average based only on sites that are mostly free of any urban heating influence. The very-rural sites could potentially be surrounded by a built-up environment at a scale smaller than the resolution of the MOD500 map, but we assume that any resulting heating would be small compared to a city.

Figure 2 Locations of the 36,869 stations in the Berkeley Earth data set



3. Station Trend Analysis

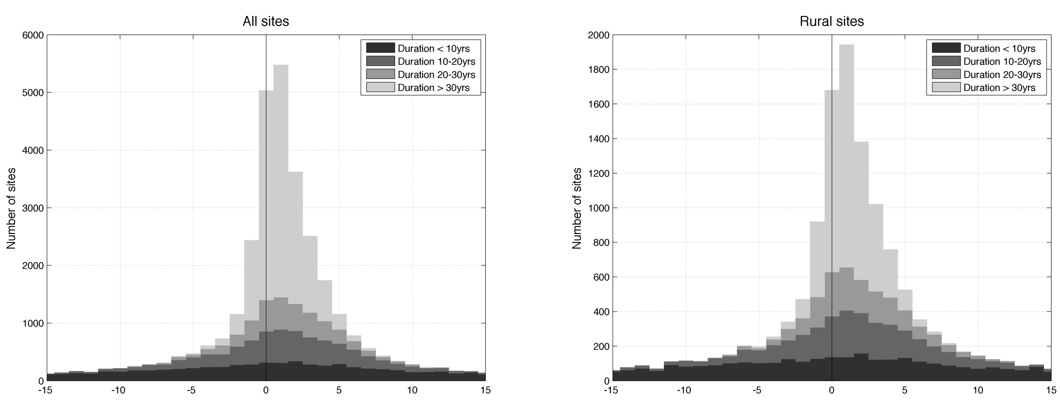
A straightforward way to gain insight into the temperature trends associated with the stations in very-rural locations is a station trend analysis. We apply a very simple procedure in which a straight line is fit (using least squares minimization) to the temperature record for each station; the slope of this line is called the temperature trend for that station. The distribution of these trends can then be examined. For the purposes of this simple analysis, we do not consider whether any individual trend is statistically significant. In fact, we expect many trends are driven primarily by statistical fluctuations and noise, but by looking at such trends in the aggregate we can yield some basic insights about the population of station time series from which they are derived. A primary limitation of the trend analysis is that it is an average over stations and time, not an average over the true land distribution of the Earth or the distribution of recording stations though time. Nevertheless, this technique has the advantage of simplicity, and it illustrates important features of the temperature record.

For the station trend analysis, we used the data set of the Berkeley Earth project consisting of the raw data for each of 36,869 sites with seasonality removed (Berkeley Earth Merged Dataset version 2 – TAVG Monthly - Non-seasonal / Quality Controlled , http://berkeleyearth.org/data/).

A histogram of the station trends is shown in Figure 3a, categorized by station record length. The distribution is broad with a width substantially larger than the mean; 65% of the slopes are positive, i.e. there are about twice as many stations that appear to warm as stations that appear to cool. The dispersion is larger in the records of short duration, but even in the stations with records longer than 30 years 24% have negative trends.

The reason the records with the shortest duration (< 10 years) have the broadest distribution is that short term variations in individual time series are typically several degrees C, so a 2 degree fluctuation during a 10 year period could yield an apparent "trend" of 20 degrees per century. There were other causes for spuriously large trends; for example, in some samples there is a gap in the data lasting for years or decades, with a large jump in the value of the average temperature when the data resumes. This is likely due to undocumented station changes and/or the reuse of an existing site identifier. Very large trends are largely non-physical and trends more extreme than ±15 ^{o}C/100yr are excluded from the histogram but not the following calculations; this excludes about 17 % of all sites but only 0.75% of sites with records longer than 10 years. To avoid the outliers unduly influencing of estimates of the center of the distributions we compare medians rather than means.

Figure 3 Temperature trends



The median trends with standard errors are given in Table 1.

Table 1. Estimates for the median trends for all and rural stations[2]

Station characteristic	Median trend in ^{o}C/100yr	
	Sites with ≥ 2 months	*Sites with >30 years*
all	0.96 ± 0.03 (n = 36858)	0.87 ± 0.02 (n = 14481)
very rural	1.10 ± 0.06 (n = 15587)	1.02 ± 0.05 (n = 4765)

The standard errors were obtained by bootstrap resampling (sampling with replacement) of all 36858 trends, calculating the median trend in each group, and using the standard deviation of the group medians to estimate the standard error in the overall median. They do not take into account the spatial correlation of the trends and hence underestimate the true uncertainty in the estimate of the median trend.

In this table we see evidence of "global warming." Using all the records there is a median warming trend of 0.96 ± 0.04 ^{o}C/100yr (2σ error). The estimated warming trend for the very-rural group is larger than that based on all records, in the opposite direction expected from urban heating. The difference observed in this simple analysis reflects that there are many sources of variation in individual station trends that contribute larger effects than any effect due to urban heating. To extract the urban heating contribution from the individual trends a careful analysis would involve modeling known sources of variation,

[2] The number of stations in each group is shown in brackets. Stated errors are 2σ uncertainty estimated from bootstrap samples.

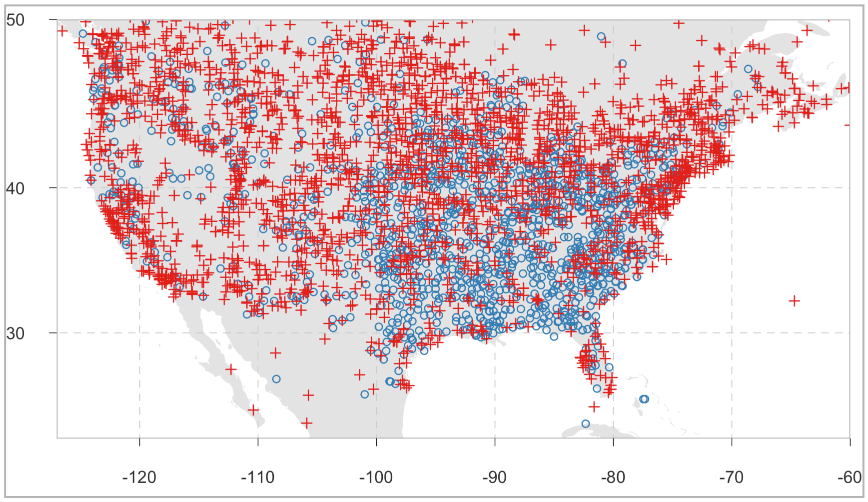
such as geographic location and measurement differences as well as accounting for spatial distribution and correlation. Since our primary interest is to evaluate the effect in the global average we do not pursue this approach further.

Although trend analysis is a very crude way to look at global temperature change, it illustrates important features of the data. The histograms show that the global warming is in some ways a subtle effect compared to the weather and instrumental noise that can affect individual stations. The distribution of trends in the station data is so broad that many simultaneous measurement sites are necessary in order to properly characterize the effect; a handful is not enough. With a full width at half max of about 5 ^{o}C per century, the trend histogram suggests that averaging one hundred independent stations would yield a 1σ trend uncertainty of about $5/\sqrt{100} = 0.5$ ^{o}C/century – just barely enough to resolve the collective temperature trend (compare to Jones (1994) who found a subset of 172 well dispersed stations gave a reasonable estimate of the global average). With over 30,000 stations, we do much better. The trend analysis also supports the view that the spurious contribution of urban heating to the global average, if present, is not a strong effect; this agrees with the conclusions in the literature that we cited previously.

The positive and negative sloped stations are intermingled, even though some patterns related to underlying climate also occur. This is seen in Figure 4, a map of the stations in the United States with at least a 70 year duration, with red + signs indicating stations that showed net warming over their record, and blue circles showing stations with net cooling. As with the world sample, the ratio of warming sites to cooling ones was in the ratio of 2:1.

Some spatial homogeneity is present, but it is nonetheless possible to find long time series with both positive and negative trends from all portions of the United States.

Figure 4. Map of stations in and near the United States

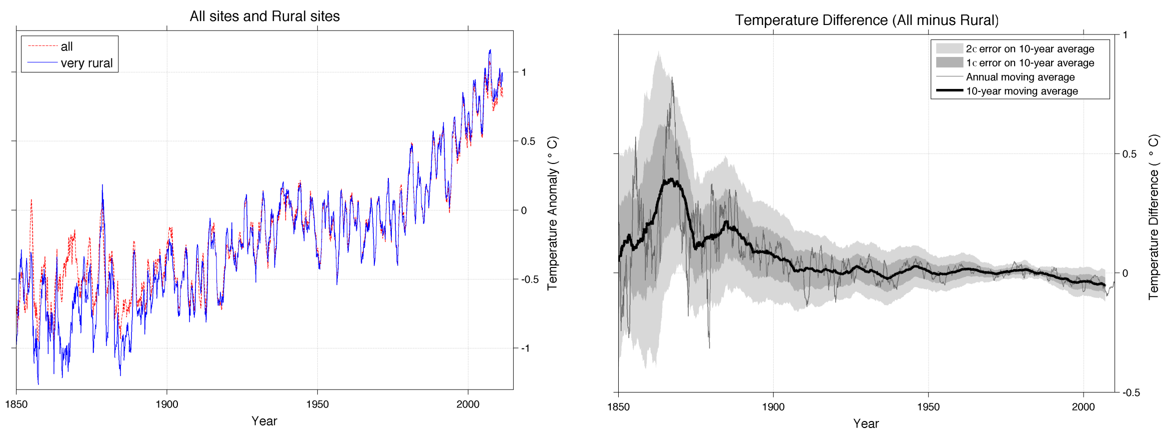


4. Berkeley Earth Surface Temperature Global Average

For a more rigorous estimate of the urban heat island effect, we performed a complete global land temperature record reconstruction using the Berkeley Earth Surface Temperature averaging methodology (Rohde et al., 2012, draft available at http://berkeleyearth.org/pdf/berkeley-earth-averaging-process.pdf). Briefly, this includes the following steps. Prior to averaging, station records are broken using metadata at times of changes in time of observation, station location, and at gaps in their records. They are also broken at times they exhibit changes in the statistical properties of their record by comparing their record to other local stations. In the averaging process stations are weighted according to their spatial correlation as well as their reliability. Uncertainties on the average are calculated that incorporate both statistical uncertainty and uncertainty due to spatial incompleteness. We evaluate the effect of very-rural station siting on the global average by applying the Berkeley Earth Surface Temperature averaging procedure to the very-rural stations. By comparing the resulting average to that obtained by using all the stations we can quantify the impact of selecting sites not subject to urbanization on the estimated average land temperature.

In the full averaging procedure sites have their weights adjusted via an iterative procedure which compares their time series to the reconstructed T_{avg}; sites that deviate substantially from the local group behavior have their weights reduced for the next iteration (see Rohde et al. (2012) for details). Thus, the influence of sites with anomalous trends, such as urban heat island effects, should be reduced by the averaging procedure even when sites with spurious warming are part of the dataset being considered. In Figure 5A we show the comparison of the temperature estimate for all the land sites (in red) with the temperature trend for the very rural land sites (blue). The difference between the two plots is shown in Figure 5B. An urban heat island bias would be expected to show itself as an upward trend in 5B; none is seen.

Figure 5. **A.** Berkeley Earth global temperature averages, normalized to zero mean for 1950-1980. **B** is the difference between the two curves in **A**.



Over the bulk of the record, the difference between the two calculations is consistent with zero within 2 standard errors (shown as the light grey area on Figure 5B, the standard errors are calculated by adding the statistical uncertainties calculated for each average, using the Jackknife method of Rohde et al. (2012), in quadrature). At later times a slight downward trend is observed, but it is not statistically significant. Over the period 1950 to 2010 (covering most of the data in Fig 3, and during which anthropogenic interference with climate is considered most acute) the temperature difference (Fig 5B) had a slope of -0.10 +- 0.24 °C/100yr (2σ error). The error on the slope includes the statistical errors on the global land average at each time point. While the point estimate is in the opposite direction to urban heating the interval is consistent with zero urban heating and the heating effect estimated by the prior groups of +0.01 to +0.1°C per century. Sensitivity to the definition of very-rural being 0.1 degree distant from an urban area was assessed by repeating the analysis in this section using distances of 10km and 25km. The resulting curves were very similar, for the period 1950 to 2010 the difference between the two curves was -0.11 +-

0.20 degrees C / century (2σ error) for very rural defined as at least 10km from an urban area (15,374 stations), and -0.11 +- 0.22 degrees C / century (2σ error) for very rural defined as at least 25km from an urban area (9,670 stations).

5. Discussion

We observe the opposite of an urban heating effect over the period 1950 to 2010, with a slope of -0.10 +- 0.24 °C/100yr (2σ error) in the Berkeley Earth global land temperature average. The confidence interval is consistent with a zero urban heating effect, and at most a small urban heating effect (less than 0.14°C/100yr, with 95% confidence) on the scale of the observed warming (1.9 ± 0.1 °C/100yr since 1950 in the land average from figure 5A).

The stations we identified as "very rural" provide good spatial coverage of the land surface of the globe and an average based solely on these stations provides a reconstruction robust to urban heating. Our results are inline with previous results on global averages despite differences in methodology. Parker (2010) concluded that the effect of urban heating on the global trends is minor, HadCRU use a bias error of 0.05 °C per century, NOAA estimate residual urban heating of 0.06 °C per century for the USA and GISS applies a correction to their data of 0.01 ^{o}C per century. All are small on the scale of global warming.

The huge effects seen in prominent locations such as Tokyo have caused concern that the T_{avg} estimates might be unduly affected by the urban heat effect. It did not have a strong effect on our estimate - which is not surprising given that urban areas are only 0.5% of the land area (according to the MOD500 map). The station slope analysis shows that there are also a large number of sites with negative trend lines. Some of these could be due to cooling effects resulting from anthropogenic changes to the landscape. For example, in an urban area if an asphalt surface is replaced by concrete, we might expect the solar absorption to decrease, leading to a net cooling effect. Rural areas could show temperature biases due to anthropogenic effects, for example, changes in irrigation.

We note that our averaging procedure uses only land temperature records. Inclusion of ocean temperatures will further decrease the influence of urban heating since it is not an ocean phenomenon. Including ocean temperatures in the Berkeley Earth reconstruction is an area of future work.

6. Acknowledgements

This work was done as part of the Berkeley Earth project, organized under the auspices of the Novim Group (www.Novim.org). We thank many organizations for their support, including the Lee and Juliet Folger Fund, the Lawrence Berkeley National Laboratory, the William K. Bowes Jr. Foundation, the Fund for Innovative Climate and Energy Research (created by Bill Gates), the Ann and Gordon Getty Foundation, by the Charles G. Koch Charitable Foundation, and three private individuals (M.D., N.G. and M.D.). More information on the Berkeley Earth project can be found at www.BerkeleyEarth.org.

7. References

Arnfield, A. J. (2003), Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island, *International Journal of Climatology*, 23, (1), 1--26.

Brohan, P., J. Kennedy, I. Harris, S. Tett, and P. Jones (2006), Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850, *J. Geophys. Res*, 111, (D12).

Chandler, T. (1976), The climate of towns, *The climate of the British Isles*, 307--329.

De Laat, A., A. Maurellis (2006), Evidence for influence of anthropogenic surface processes on lower tropospheric and surface temperature trends, *International Journal of Climatology*, 26, (7), 897--913.

Fujibe, F. (2011), Urban warming in Japanese cities and its relation to climate change monitoring, *International Journal of Climatology*, 31, (2).

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

Howard, L. (1833), The climate of London, deduced from meteorological observations, 2.

Jones, P. D. (1994), Hemispheric Surface Air Temperature Variations: A Reanalysis and an Update to 1993, *Journal of Climate*, 7, (11), 1794--1802.

Jones, P., S. Raper and T. Wigley (1986b), Southern Hemisphere surface air temperature variations: 1851--1984, *J. Appl. Meteorol.;(United States)*, 25, (9).

Jones, P., S. Raper, P. Kelly, T. Wigley, and R. Bradley (1986a), Northern hemisphere surface air temperature variations 1851-1984, *Journal of climate and applied meteorology*, 25, (2), 161--179.

Jones, P., P. Y. Groisman, M. Coughlan, N. Plummer, W. Wang, and T. Karl (1990), Assessment of urbanization effects in time series of surface air temperature over land, *Nature*, 347, (6289), 169--172.

Karl, T. R., H. F. Diaz and G. Kukla (1988), Urbanization: its detection and effect in the United States climate record, *Journal of Climate*, 1, (11), 1099--1123.

McKitrick, R., P. J. Michaels (2004), A test of corrections for extraneous signals in gridded surface temperature data, *Climate Research*, 26, (2), 159--173.

McKitrick, R. R., N. Nierenberg (2010), Socioeconomic Patterns in Climate Data, *Journal of Economic and Social Measurement*, 35, (3,4), 149-175.

McKitrick, R. R., P. J. Michaels (2007), Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data, *J. Geophys. Res.*, 112, (D24).

Menne, M. J., C. N. Williams Jr (2009), Homogenization of temperature series via pairwise comparisons, *Journal of Climate*, 22, (7), 1700--1717.

Menne, M. J., C. N. Williams Jr and R. S. Vose (2009), The US Historical Climatology Network monthly temperature data, version 2, *Bulletin of the American Meteorological Society*, 90, (7), 993--1007.

Oke, T. (1979), Review of Urban Climatology, 1973-1976.

Oke, T. (1974), Review of Urban Climatology, 1968-1973.

Oke, T. (1982), The energetic basis of the urban heat island, *Quarterly Journal of the Royal Meteorological Society*, 108, (455), 1--24.

Parker, D. E. (2004), Large-scale warming is not urban, *Nature*, 432, (7015), 290--290.

Parker, D. E. (2010), Urban heat island effects on estimates of observed climate change, *Wiley Interdisciplinary Reviews: Climate Change*, 1, (1), 123--133.

Peterson, T. C. (2003), Assessment of urban versus rural in situ surface temperatures in the contiguous United States: No difference found, *Journal of Climate*, 16, (18), 2941--2959.

Peterson, T. C., K. P. Gallo, J. Lawrimore, T. W. Owen, A. Huang, and D. A. McKittrick (1999), Global rural temperature trends, *Geophysical Research Letters*, 26, (3), 329--332.

Potere, D., A. Schneider, S. Angel, and D. L. Civco (2009), Mapping urban areas on a global scale: which of the eight maps now available is more accurate?, *International Journal of Remote Sensing*, 30, (24), 6531--6558.

Rohde, R., D. Brillinger, J. Curry, D. Groom, R. Jacobsen, E. Muller, R. Muller, S. Perlmutter, A. Rosenfeld, C. Wickham, and J. Wurtele (2012), Berkeley Earth Temperature Averaging Process.

Schmidt, G. A. (2009), Spurious correlations between recent warming and indices of local economic activity, *International Journal of Climatology*, 29, (14), 2041--2048.

Schneider, A., M. Friedl and D. Potere (2009), A new map of global urban extent from

MODIS satellite data, *Environmental Research Letters*, 4.

Schneider, A., M. Friedl and D. Potere (2010), Mapping global urban areas using MODIS 500-m data: New methods and datasets based on urban ecoregions', *Remote Sensing of Environment*.

8. Figure Captions

Figure 1 Annual running mean of monthly temperatures at Tokyo compared to a global land average for 1900-2010. (Tokyo station id: wmo_47662).

Figure 2 Locations of the 36,869 stations in the Berkeley Earth data set. (a) 15,594 stations classified as very-rural, at least 0.1° from an urban area in the MOD500 map (Schneider et al. 2009, 2010). (b) Locations of the other 21,271 stations.

Figure 3 Temperature trends. A histogram of the trends is shown in (a) for all land stations in the Berkeley Earth data set of 36,869 records, and (b) only rural stations, defined as those that are at least 0.1 degrees in latitude and longitude from a MOD500 urban region. The x-axis limits are chosen to include the central 80% of trends in (a).

Figure 4. Map of stations in and near the United States with at least 70 years of measurements; red "+" stations are those with positive trends and blue "o" stations are those with negative trends.

Figure 5. A. Berkeley Earth global temperature averages, normalized to zero mean for 1950-1980. The dashed (red) estimate is based on all sites; the solid (blue) estimate is based on the very rural sites (those more than 0.1 degrees distant from a MOD500 urban region). B is the difference between the two curves in A. The thin line shows a one-year running average; the thicker line shows the 10-year running average. The dark grey area shows the

standard error on the 10-year running average, the light grey twice the standard error on the 10-year running average. The standard errors are calculated by adding the statistical uncertainties calculated for each average, using the Jackknife method of Rohde et al. (2012), in quadrature.