

Urban Heat Island Effects in U.S. Summer Surface Temperature Data, 1895-2023



Roy W. Spencer,^a John R. Christy,^{a,b} William D. Braswell^a

^a *Earth System Science Center, The University of Alabama in Huntsville, Huntsville, Alabama*

^b *Alabama Office of the State Climatologist, Huntsville, Alabama.*

Corresponding author: Roy W. Spencer, roy.spencer@nsstc.uah.edu

Early Online Release: This preliminary version has been accepted for publication in *Journal of Applied Meteorology and Climatology*, may be fully cited, and has been assigned DOI 10.1175/JAMC-D-23-0199.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

© 2025 American Meteorological Society. This is an Author Accepted Manuscript distributed under the terms of the default AMS reuse license. For information regarding reuse and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

ABSTRACT

A novel method is described for quantifying average urban heat island (UHI) warming since 1895 in contiguous U.S. (CONUS) summer air temperature data. The method quantifies the sensitivity of Global Historical Climatology Network (GHCN) station raw temperature to station-centered population density (PD). Specifically, closely spaced station pair differences in monthly raw (non-homogenized) T_{AVG} (the average of daily maximum and minimum temperature) and PD are sorted by station pair average PD into six PD classes, and linear regression estimates of the temperature sensitivity to population density change (dT_{AVG}/dPD) are made for each class for historical periods ranging from 1 to 21 years in length. Every one of the resulting six sensitivity relationships in each of 22 historical periods from 1880 to 2020 are found to be positive, and their magnitudes allow construction of station-average urban heat island temperature (T_{UHI}) curves as a function of population density. When applied to the history of population changes at each CONUS station location (1895-2023) and grouped into four categories of station population density, the resulting T_{UHI} warming trends range from 8% of observed T_{AVG} warming for the most rural category of stations to about 65% of observed warming for suburban and urban categories. Across all stations the UHI warming amounts to 22% of the observed raw GHCN warming trend, (+0.016 versus +0.072 °C decade⁻¹). The method provides an independent way to quantify station-average UHI warming over time.

1. Introduction

The urban heat island (UHI, Balchin & Pye, 1947) is a ubiquitous feature of human settlements, first described in *The Climate of London*, by Luke Howard in 1833 (Mills, 2008). The UHI warming effect on air temperatures is maximized at night and is mostly due to the replacement of vegetation and aerated soils with buildings and impervious pavement (Arnfield, 2003). These changes alter the local day-night energy budget of solar energy gain and energy loss by infrared radiation, evaporation, and convective air currents compared to wilderness conditions (Oke, 1982). Since the resulting warming affects the health and comfort of millions of city-dwelling people around the world, the UHI enhancement of global and regional warming trends is of increasing concern to the public, healthcare workers, urban planners, and policymakers (Estoque et al., 2020; Tong et al., 2021; Hsu et al., 2021; Zhang et al., 2023; Chen et al., 2023).

Since most surface air temperature measurements are made in or near human settlements, and most of those settlements have grown over time, it is reasonable to address whether global warming trends calculated for land areas have been spuriously inflated. Unfortunately, quantifying the UHI effect on air temperature is complicated by a wide variety of contributing factors and how they have changed over time: population density, building height and density, roads and parking lots, active sources of waste heat, and settlement size and heterogeneity just to name a few (Hausfather et al., 2013; Peterson & Owen, 2005; Gallo & Owen, 2002; Stewart and Oke, 2012). It should be noted that UHI effects are much stronger in daily minimum (T_{MIN}) than maximum (T_{MAX}) temperatures, by about a factor of three (e.g. Hamdi & Van deVyver, 2011). Homogenization procedures are used in the adjusted version of the GHCN dataset (Menne and Williams, 2009; Menne et al., 2012) which involve correcting for breakpoints in a station's temperature time series relative to surrounding stations. To the extent that UHI warming is not uniform across all nearby stations, this is believed to remove spurious temperature trends from UHI in the adjusted (homogenized) dataset.

With a few exceptions (e.g. Mitchell, 1953; Kukla et al., 1986; Scafetta, 2021) most efforts to determine whether the UHI has spuriously inflated land warming trends have concluded it has little effect (Wickham et al., 2013; Hansen et al., 2010; Parker, 2010; Jones et al., 2008; Parker, 2006; Peterson & Owen, 2005; Peterson, 2003; Peterson et al., 1999; Gallo et al., 1999; Karl et

al., 1988). Hausfather et al. (2013) estimated the UHI effect in both the raw and homogenized versions of the GHCN dataset and found 14-21% of the rise in raw (unadjusted) minimum temperatures during 1895-2010 could be traced to urbanization. Others have claimed that homogenization procedures are not well behaved and have produced spurious effects in the temperature record (Soon et al., 2018; O'Neill et al., 2022; Katata et al., 2023). Given that homogenization adjustments correct for any source of step changes (including UHI) in a station's temperature time series relative to neighboring stations, we believe it is worthwhile to isolate the UHI effect in the raw temperature data using an alternative method.

We quantify the UHI signal in the spatial domain using GHCN summer monthly (June, July, August) raw T_{AVG} (the average of daily maximum and minimum temperatures) and two gridded global population density datasets. Since this is an exploratory effort meant to establish a methodology for quantifying UHI effects, we restrict our analysis to the summer months, when UHI effects contribute to the highest urban temperatures. We are guided by the seminal work of Summers (1964) and Oke (1973) which showed a strong nonlinearity in the relationship between UHI warming and settlement population. This nonlinearity, which we will demonstrate with GHCN data, has important implications when comparing urban to rural temperature trends. For example, Oke's reported 4th root power law relationship between nighttime temperature and population for cities and towns in Canada, the United States, and Europe would result in more warming from a rural population density increase from 1 to 10 persons km^{-2} than would an urban increase from 1,000 to over 1,700 persons km^{-2} . This suggests the common assumption that air temperatures in rural locations are unaffected by urbanization is not strictly true. If rural locations have experienced some level of UHI warming, data adjustment methods that assume they can be used as a baseline for computing urbanization effects at non-rural locations might underestimate those UHI effects.

Given this nonlinearity, the UHI problem would benefit from new methods for quantifying warming that do not depend upon the assumption that rural stations are unaffected by UHI. We describe such a method for quantifying the average UHI effect across thousands of stations representing a full range of population densities from wilderness to inner-city. Our reconstruction of the functional relationship between population density and UHI warming (T_{UHI}) uses linear regression of closely spaced paired-station differences in temperature (ΔT_{AVG}) against

differences in population density (ΔPD). The novel feature of our analysis is the sorting of these many thousands of station pairs into six classes of paired-station average population density for linear regression. The regression coefficients represent the local slopes of the T_{AVG} vs. PD curve which then allows the nonlinear functional relationship between station-average T_{UHI} warming and population density to be quantified.

2. Data and Methods

The GHCN global dataset includes over 27,000 stations with various lengths of record operating since the late 1800s, with nearly half of those stations located in the U.S. In the time domain it is difficult to disentangle UHI effects from the global warming signal because the signals are often similar: both warming of the climate system and urbanization occurs slowly, over many decades. But to the extent such changes tend to be sporadic in time and space, homogenization techniques have been developed (Menne and Williams, 2009; Menne et al., 2012) which assume urbanization effects with time are not spatially uniform across neighboring stations and then adjust for relative temporal step changes in temperature. This constitutes the “adjusted” (homogenized) version of the GHCN dataset. It is beyond the scope of this paper to determine how much (if any) urban warming remains in the adjusted (homogenized) GHCN data; here we will compute UHI effects from the raw (unadjusted) version of the dataset.

In the spatial domain it is widely recognized — even by the lay person in their daily commute — that urban environments are warmer than rural areas. In the United States, these urban areas have grown substantially since the late 1800s when temperature datasets started for many stations. We estimate the average spatial urban warming in the raw summer GHCN data across many stations using 2-station paired differences in temperature and population density to quantify how UHI warming changes with population (dT_{AVG}/dPD), then we apply those relationships to individual station population changes over time (dPD/dt). The technique can, of course, be applied to other seasons besides summer. The extension of the UHI signals quantified in the spatial domain to UHI warming in the time domain is accomplished with Eq. 1,

$$dT_{UHI}/dt = [dT_{AVG}/dPD][dPD/dt], \quad (1)$$

where T_{UHI} is the urban heat island component of temperature, t is time, and PD is population density in persons km^{-2} . Note there are two parts to Eq. 1, the first being the change in

temperature with population density, and the second being the change in population density with time. The first calculation is the most involved, with a number of required steps, while the second is a relatively straightforward calculation from the population density dataset. Also note that once a time history of dT_{UHI}/dt is computed with Eq. 1, then time integration over the period of the instrumental record leads to a time series of just the UHI component of temperature, T_{UHI} .

We quantify spatial temperature dependence on population density dT_{AVG}/dPD by computing the average 2-station temperature and population density differences, and binning those averages into six population density bins. As a simple qualitative example, urban locations average warmer than suburban locations, and suburban warmer than rural, but with different magnitudes. Once the dT_{AVG}/dPD values are computed, curves of T_{UHI} vs. PD can be constructed. We use six classes of population density ranging from heavily populated to near wilderness, in 22 historical periods between 1880 and 2020. We tried as many as 20+ classes of population density, but the results did not substantially depend upon this choice, so we settled on six as the minimum number to mostly capture the nonlinear relationship between T_{AVG} and PD . We admit this methodology is heuristic, and a more objectively-chosen class definition scheme might produce somewhat better results. We have found that still fewer classes might be necessary if the method is extended to non-U.S. regions where there are fewer stations in sufficiently close proximity to one other to allow meaningful regression analysis, otherwise the regression coefficients become too noisy to construct meaningful curves of T_{UHI} vs PD .

Specifically, we performed both ordinary least squares (OLS) linear regressions and Deming regressions (Deming, 1943) between 2-station differences in monthly average air temperature for the historical periods in Table 2 (covering 1880-2020) from the GHCN monthly Version 4 raw data (Menne et al., 2018) and 1/12th degree grid ($\sim 10 \times 10$ km) population density from the HYDE 3.3 dataset (Goldewijk et al., 2017) to estimate dT_{AVG}/dPD regression coefficients. Regressions are done separately for different classes of 2-station average PD , and a zero intercept is not assumed in the regressions since we are estimating the local slopes along the T vs PD curve, which is nonlinear. The proportion of stations assigned to each of the six 2-station average PD classes (33%, 30%, 20%, 12%, 4.5% and 0.5%) was meant to capture the functional nonlinearity across a full range of PD values while keeping the regression t-statistics relatively high in all classes. These proportions were used for all subsequent regressions. As discussed below, Deming regression is one method to adjust for low biases in the regression coefficient

that result from errors in the independent variable (in this case, population density). The HYDE population data is reported every ten years between 1880 and 1950, and then every year until 2023, and for each station location we interpolate those reported values in the decadal data to yearly time resolution using Hermite polynomial interpolation. We also used the Global Human Settlement (GHS-POP, Freire et al., 2016) population dataset, available every five years from 1975 to 2020, in order to compare to the HYDE PD data for the purpose of estimating PD errors necessary for the Deming regressions.

Stations included in the matchups had to be separated by not more than 120 km in distance and 250 m in elevation. Larger separations would lead to more station pairs to include in the regressions, but at the expense of adding more noise since weather conditions (rain, clouds, synoptic temperature variations, elevation effects) differ more at greater separations. Smaller separations lead to less weather noise, but fewer station pairs. The choice of 120 km and 250 m was made after some experimentation, but should not be considered to be optimized. We then apply these dT_{AVG}/dPD relationships to each station's PD history (dPD/dt) to estimate the UHI effect on temperature over time (Eq. 1), which can be compared to temperature trends calculated from the GHCN data.

3. Results

a. Computations of dT_{AVG}/dPD

As an example, from one of the 22 historical periods, the resulting dT_{AVG}/dPD regression statistics for 1975 are shown in Table 1. Note that each of the six classes of 2-station average population density results in a positive regression coefficient, in decreasing magnitude as PD increases. This is evidence of the nonlinear relationship between population and temperature described by Oke (1973) and others.

PD Class	Number of 2-Station Pairs (N) /Fraction in Class	Correl. Coeff.	OLS Regression Coeff. (dT_{AVG}/dPD)	Deming Regression Coeff. (dT_{AVG}/dPD)	2-station Average PD & Range (persons km ⁻²)	OLS t-statistic
1	71,175/33.0%	0.110	0.00554	0.00679	17.4 (0.005 to 41)	29.5
2	64,704/30.0%	0.177	0.00215	0.00254	76.3 (41 to 125)	45.9
3	43,136/20.0%	0.226	0.00118	0.00132	201.5 (125 to 314)	48.3
4	25,882/12.0%	0.187	0.00050	0.00056	497.0 (314 to 859)	30.7
5	9,706/4.5%	0.345	0.00034	0.00037	1,496.5 (859 to 3,798)	36.2
6	1,076/0.5%	0.734	0.00012	0.00012	6,002.3 (3,798 to 12,962)	35.4

Table 1. Example OLS and Deming linear regression results between CONUS 2-station differences in raw GHCN temperature and population density for June, July, and August in just one of the 22 historical periods, 1975. Six classes of 2-station average population density regression results including number of station pairs N, Pearson correlation coefficient, OLS and Deming regression coefficient estimates of dT_{AVG}/dPD , and t-statistic of the OLS regression. The regression error statistics (t-statistics) in the last column are very much larger than 2 (the commonly accepted value for statistical significance), but are likely overestimated since the 2-station pairs are not all independent, often involving the same stations but at different times and in different combinations.

The Deming regression results in Table 1 require some explanation. As can be seen in Table 1, the correlation coefficients for the most rural population density classes are quite low, below 0.2 for the first two classes. This naturally leads to concerns regarding the usefulness of the resulting regression coefficients, despite the large t-statistic values. For this reason, we also include the Deming regression results, which is one of several “errors in variables” methods for regressing particularly noisy data. As was originally addressed by Spearman (1904) the quantified relationship (regression slope) between two noisy variables will be biased toward zero (attenuated) to the extent that there is noise in the independent variable, in this case population density. In contrast, for OLS regression, increasing noise in the dependent variable (temperature) will lead to increased uncertainty in the regression coefficient, but the coefficient estimate will remain unbiased as long as the independent variable (PD) is noise-free. “Noise” in this case can

be either measurement errors in the variables, or sources of noise in either of the variables unrelated to UHI. “Errors in variables” methods for correcting for the attenuation of the regression coefficient have been addressed by Frost and Thompson (2000), and the issue has also been described in Fuller (1987) and Hutcheon et al. (2010). More recently, McKittrick (2023) addressed the possibility that errors in variables techniques can over-correct for regression attenuation if not carefully matched to the characteristics of the dataset in question.

We employed Deming regression (Deming, 1943) to obtain unbiased estimates of the dT_{AVG}/dPD regression coefficients, also listed in Table 1, which requires error estimates for both the 2-station differences in T_{AVG} and PD . The T_{AVG} difference errors were assumed to be the standard deviation of the OLS regression coefficient errors. For the PD difference error, we use the standard deviation of the differences in the independent GHS-POP and HYDE station PD estimates, at 10 km resolution. The resulting PD errors averaged from 68% of the most rural class average PD to 27% of the most urban class average PD .

The resulting regression coefficients in Table 1 represent the local slopes of the T vs. PD curves, which as previously mentioned other researchers have determined to be nonlinear. A comparison of the OLS-based versus Deming-based results is shown for summer of 1975 in Fig. 1, where (a) shows the regression coefficients as a function of 2-station average population density, and (b) shows the integration of the power law fits to those data across PD using the trapezoidal rule.

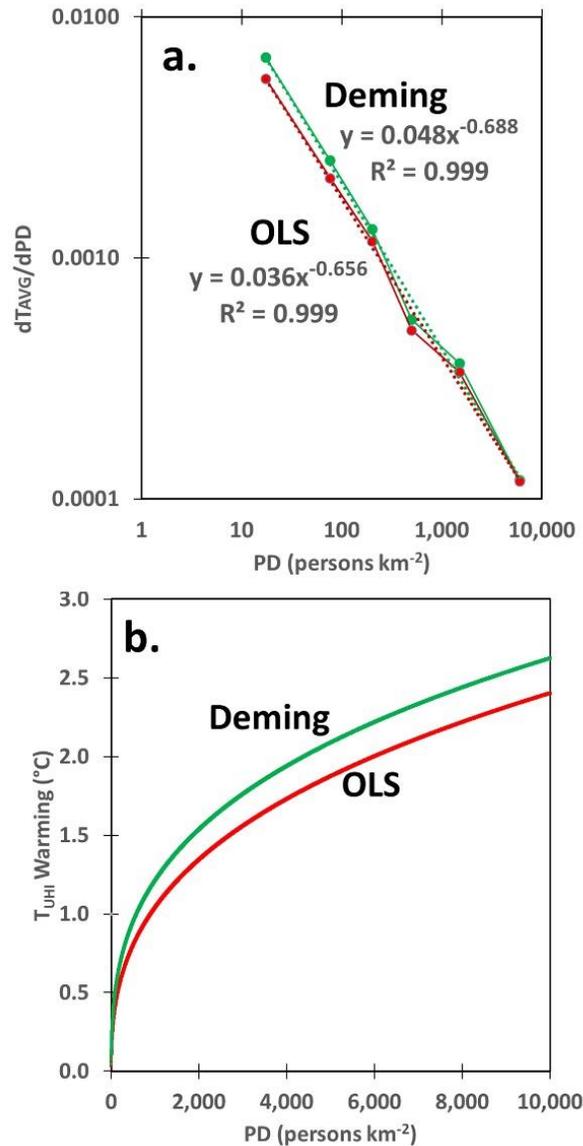


Fig. 1. (a) dT_{AVG}/dPD regression coefficients for the six 2-station average population density classes in summer 1975 U.S. GHCN raw temperature data and their power-law regression fits (dotted lines) from Table 1, and (b) the resulting UHI warming curves from integrating the power law fits in (a) from zero to PD values from 1 to 10,000 persons km^{-2} using the trapezoidal rule.

As seen in the log-log plot in Fig 1a, a power law relationship fits the regression coefficients very well for both OLS and Deming regressions, with 99.9% explained variance using 1975 data. This explained variance represents how well a power law relationship describes the increase in warming rate with population density across six PD classes,

$$dT_{\text{AVG}}/dPD = c_1PD^{c_2} \quad (2)$$

The UHI curves in Fig. 1b are the result of integrating the power law relationship from $PD=0$ to $PD=10,000$. This was done by computing dT_{AVG}/dPD at every integer value from 1 to 10,000 using Eq. 2, linearly extrapolating a value at $PD = 0$ from the values at $PD = 1$ and $PD = 2$, and using the trapezoidal rule. In this case (1975 data) the Deming coefficients are larger than the OLS-based UHI coefficients in Fig. 1b, ranging from 35% larger at $PD = 1$ person km^{-2} to 9% larger at 10,000 persons km^{-2} . This Deming coefficient increase over OLS results is due to the estimated errors in the population density data.

The Deming regression results for all historical periods are shown in Table 2. The historical periods have differing lengths because of the large increase in number of available U.S. stations since the late 1800s. In order to have sufficient sample sizes to do regressions, the periods ranged from 21 years (1880-1900) to only 1 year since 1975.

Period	# Years	# 2-stn matchups	c1	c2	Expl. Variance
1880-1900	21	253,700	0.039857	-0.607542	96.3%
1901-1910	10	514,315	0.034572	-0.584998	99.5%
1911-1920	10	668,720	0.062109	-0.744012	99.9%
1921-1930	10	736,305	0.049599	-0.690120	99.5%
1931-1940	10	936,498	0.055038	-0.705208	99.9%
1941-1945	5	587,211	0.063096	-0.729402	99.9%
1946-1950	5	694,488	0.078626	-0.787744	99.5%
1951-1955	5	919,057	0.043952	-0.658996	99.9%
1957-1959	3	608,627	0.049027	-0.690737	99.9%
1962-1964	3	665,355	0.042961	-0.653218	99.3%
1967-1969	3	693,403	0.036253	-0.615864	98.4%
1972-1974	3	650,879	0.036753	-0.620486	98.5%
1975	1	215,679	0.048437	-0.687670	99.9%
1980	1	193,373	0.029092	-0.577355	98.6%
1985	1	178,725	0.037077	-0.646729	99.1%
1990	1	194,725	0.031589	-0.620783	98.4%
1995	1	188,106	0.035675	-0.635057	99.3%
2000	1	227,403	0.033121	-0.650182	98.6%
2005	1	279,171	0.029218	-0.628833	98.6%
2010	1	299,496	0.042056	-0.656549	99.6%
2015	1	255,165	0.035315	-0.638865	99.1%
2020	1	218,201	0.036626	-0.661953	93.3%

Table 2. Power law coefficients (c_1 , c_2) for 22 historical periods relating 2-station matchup Deming regression coefficients dT_{AVG}/dPD vs. 2-station average PD . These are the power law equations used to estimate UHI warming at individual GHCN stations as a function of station population density.

b. Year-over-year estimation of temperature trends

We desire to compute UHI effects only for those locations and times that have GHCN data. Unfortunately, many stations do not have continuous records of data from, say, 1880 to the present. To apply the spatial information on dT_{AVG}/dPD derived from regression to the time domain to evaluate UHI temperature trends at GHCN stations we use year-over-year (YOY) changes in temperature and population density from all GHCN stations having at least two sequential years of temperature data. It should be obvious that for stations having a long period of record with no missing months of data, the resulting temperature trends (and interannual variability) will be identical using a YOY method or using the traditional multi-year annual cycle and anomaly calculation method. The advantage of using a YOY approach is that even stations with very short periods of operation can be included without the difficulty of determining an average annual cycle when very few years of data (possibly only two) are available from a station. The disadvantage is that some station data will be excluded by the YOY method when there are not two sequential years of data. An extreme (and unrealistic) example would be a station that only operated every other year from 1880 to the present day. In this case, none of the station data would be included in YOY calculations, while a traditional calculation of an annual cycle and anomalies (departures from the annual cycle) would include all the data, and anomalies would be computed for all years having data.

To demonstrate the utility of the YOY method with observed data we compute contiguous U.S. area average summer (JJA) temperature variations between 1895 and 2023 using the adjusted (homogenized) GHCN station dataset, using all stations' year-over-year temperature changes binned on a 1 deg. (~100 km) grid and then area-averaged. We use the homogenized (rather than raw) station data to allow comparison with NOAA's spatially-analyzed homogenized data, thus establishing the validity of the YOY method. The result (Fig. 2) is extremely close to NOAA's NClmDiv dataset calculation for the contiguous U.S. (Vose et al., 2014), which averages station temperatures on a much finer grid (approximately 5 km), then averages those to the 344 U.S. climate divisions, then to the 48 contiguous states, then weights

the state-level averages based upon the relative sizes of the states. This shows that we can use the YOY method for our calculation of UHI time series and temperature trends using all available stations having two or more sequential years of data.

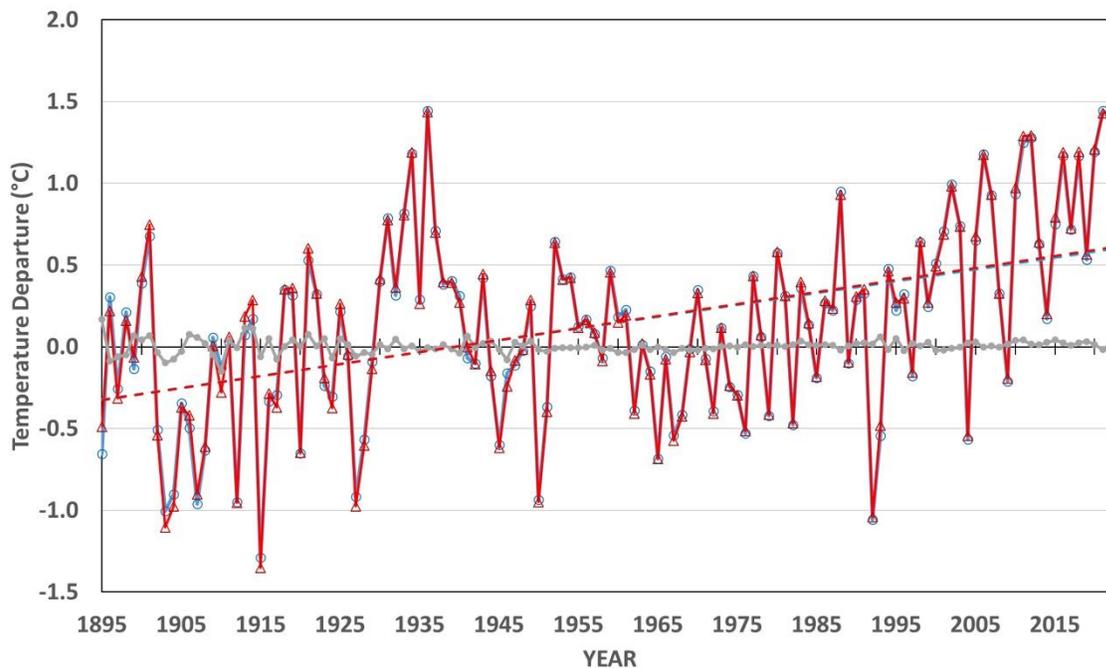


Fig. 2. Year-over-year calculation of contiguous U.S. summer T_{AVG} variations, 1895 to 2023, using the GHCN adjusted station dataset area averaged on a 1 deg. grid (red) versus the NClmDiv anomalies provided by NOAA (blue). Adjustments of coastal and border 1 deg. grid cells was made for partial cell coverage by large water bodies, Canada, and Mexico. The difference between the two datasets is shown in gray. The temperature trends (dashed lines) are nearly identical, the correlation between the YOY and NClmDiv time series is 0.997 (0.999 since 1950), and the standard deviation of the monthly differences is 0.04 deg. C (0.02 deg. C since 1950).

c. Population growth around GHCN stations since 1880 (dPD/dt)

We next document the population growth averaged across stations in different population density categories. That substantial population growth at GHCN temperature monitoring stations has occurred between 1880 and 2023 is demonstrated in Fig. 3 for four categories of station population density. The human settlement classification of the four PD categories in Fig. 3 is not well-defined in the literature, and we have chosen to label them as rural, peri-rural, suburban, and urban. In these plots an individual station can change PD category as its population grows

(or shrinks), and the station mix generally changes over time. These are averages across all stations meeting the PD criteria at different points in time, no matter where in the U.S. they are located.

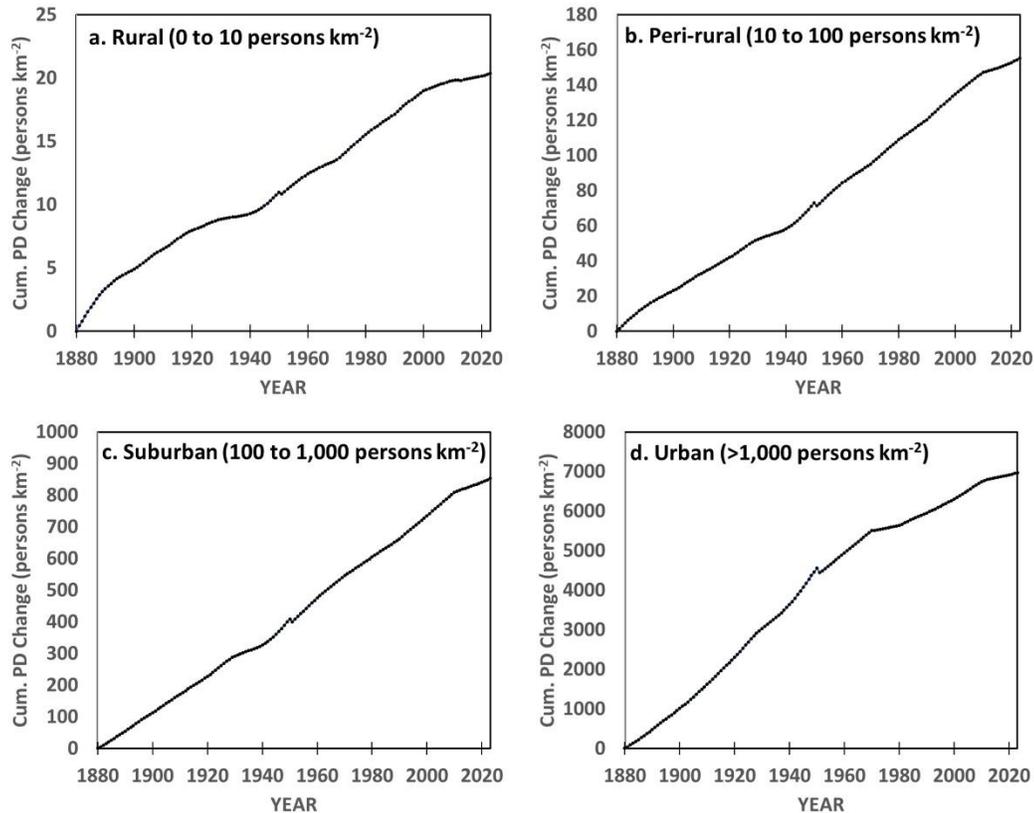


Fig. 3. Cumulative growth in population density 1880-2023 at U.S. temperature monitoring stations in four categories of station urbanization, calculated by summing the average year-on-year increases in HYDE 3.3 dataset population density ($\Delta PD/\Delta t$) at individual GHCN stations having at least two years of record for station PD of (a) 0 to 10, (b) 10 to 100, (c) 100 to 1,000, and (d) greater than 1,000 persons km^{-2} .

The cumulative average YOY growth in station population density is evident for all four population categories ranging from rural (0 to 10 persons km^{-2}) to urban (>1,000 persons km^{-2}). Therefore, we expect UHI-related warming to occur in the GHCN raw dataset in all classes of population density.

d. Dependence of UHI warming on population density in different historical periods

The regression estimates of dT_{AVG}/dPD using GHCN raw data from 1975 were listed in Table 1, and the resulting power law fits to the regression coefficients in Fig. 1a when integrated across population density lead to the urban heat island warming curves in Fig. 1b for 1975. This process was repeated for the 21 additional historical periods shown in Table 2. As in Fig. 1a, the relationships between dT_{AVG}/dPD and PD were found to be approximately linear in a log-log plot, which supports a power law relationship in every historical period as evidenced by the high explained variances in Table 2. It is those power law equation fits to the Deming regression coefficients which are then used to calculate dT_{AVG}/dPD for each station's population density value (as in Fig. 1b) over time, depending upon what historical period is being addressed.

Application of the power law equation coefficients in Table 2 to GHCN station population densities from the HYDE PD data in the 22 historical periods from 1880 to 2020 results in the examples of net UHI warming as a function of population density shown in Fig. 4, with results from OLS and Deming regression shown separately. Importantly, these are not estimates of how UHI affects temperature over time at individual stations; they represent what a hypothetical station with no change in population density would indicate for a UHI effect on temperature over time. Note there is a general tendency for the UHI effect to weaken somewhat over the last 120 years (given the same population density). We speculate this is due to improved siting of thermometers over time in terms of the microclimate environment of the stations, an effect which is not captured by our use of population density at 10 km spatial resolution.

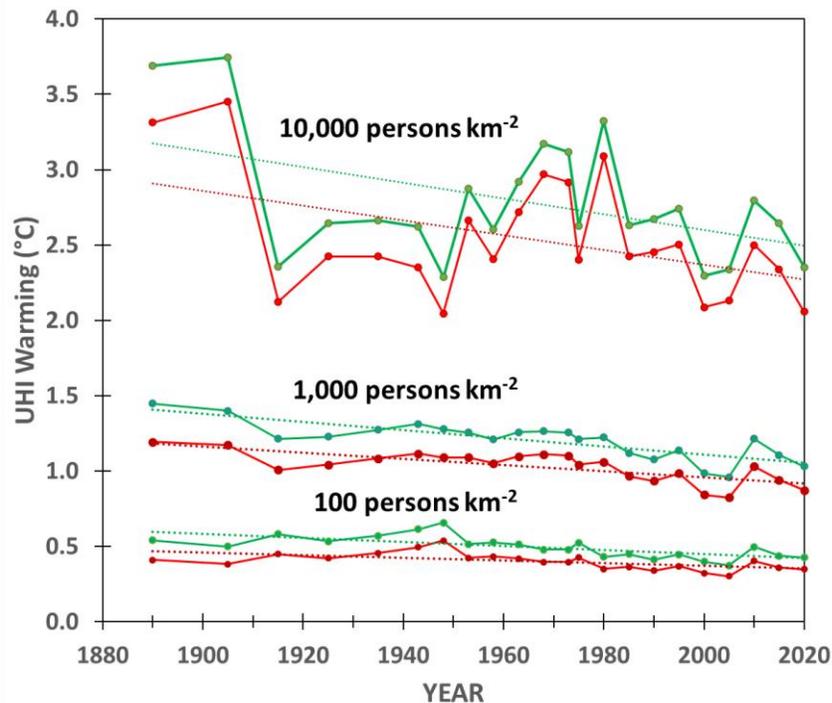


Fig. 4. Historical differences in the regression-based power law estimates of CONUS summertime UHI warming for three specific population densities ranging from peri-rural to urban center, for the historical periods listed in Table 2. The upper (green) curves are from Deming regression estimates of dT_{AVG}/dPD while the lower (red) curves are from OLS regressions. Note that very few stations have population densities approaching $10,000 \text{ persons km}^{-2}$, and so those plotted values are largely extrapolations of the power-law relationships from lower PD station values and should be considered uncertain.

e. Cumulative T_{UHI} warming in four PD classes, 1880-2020

Using the power law fits to the dT_{AVG}/dPD data in Table 2 and applying them to the year-on-year changes in PD at all CONUS GHCN stations, we next compute the cumulative T_{UHI} effect since 1880 using Eq. 1 for four separate categories of station population density (Fig. 5).

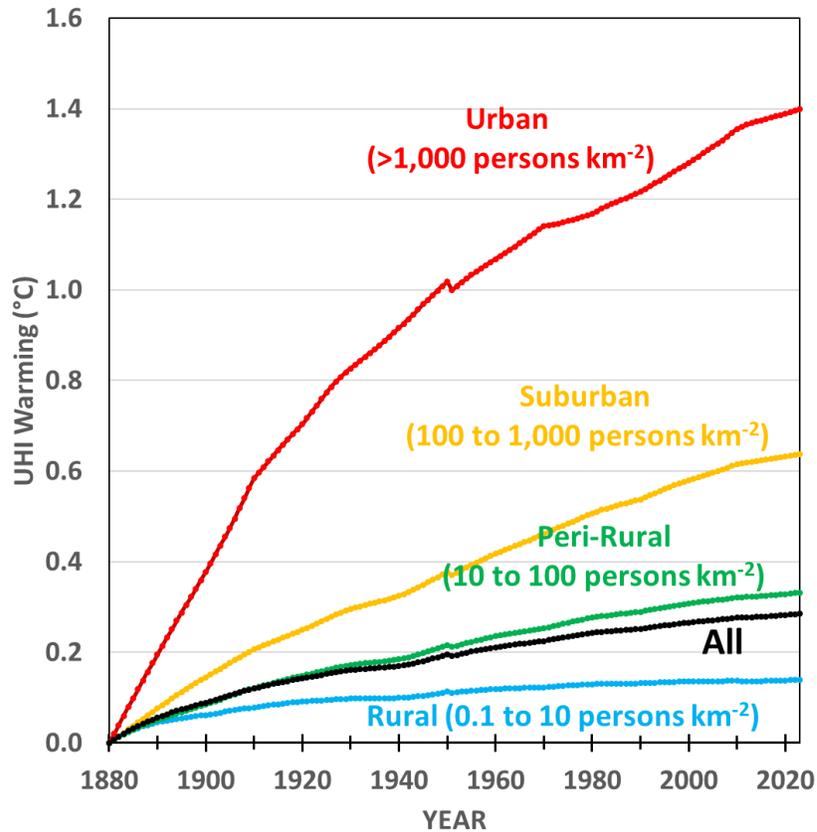


Fig. 5. Cumulative year-over-year area-average UHI warming by CONUS station population density class from dT_{AVG}/dPD calculated then applied to each station's population density history. The population density class can change for stations as they urbanize.

We see that T_{UHI} warming was stronger in the early years than in later years. The total T_{UHI} warming between 1880 and 2023 ranges from 0.14 deg. C for the least populated class to 1.4 deg. C for the most densely populated class of station, and 0.29 deg. C averaged across all GHCN stations. The “All” curve is close to the 10-100 PD class because the GHCN station mix is dominated by rural and peri-rural stations.

These T_{UHI} estimates can be compared to the station year-on-year observed temperature changes from both raw and adjusted T_{AVG} data averaged into 1 deg. grids (Table 3). It should be remembered that the different categories of station population density come from somewhat different geographic areas, with the eastern U.S. being over-represented in the high population density stations, and the western U.S. over-represented in the low population density station category. Because so few stations existed prior to 1895, we begin these comparisons in that year.

<i>PD</i> (persons km ⁻²)	1895-2023 T_{AVG} trend (°C decade ⁻¹) Raw (Adjusted)	1895-2023 T_{UHI} trend (°C decade ⁻¹)	$[T_{UHI} \text{ trend}]/[T_{AVG} \text{ trend}]$ Raw (Adjusted)
0.1 to 10	+0.075 (+0.067)	+0.006	8.2% (7.0%)
10 to 100	+0.049 (+0.056)	+0.020	41.7% (34.7%)
100 to 1,000	+0.065 (+0.083)	+0.041	63.7% (57.2%)
>1,000	+0.117 (+0.205)	+0.078	66.9% (56.8%)
All	+0.072 (+0.073)	+0.016	22.1% (21.8%)

Table 3. CONUS raw and adjusted T_{AVG} trends, and T_{UHI} trends, 1895-2023, for four station population density categories. These four categories do not represent the same geographic areas due to unequal representation of different areas by different station population densities.

For the most rural class of station, Fig. 6 shows that the 1895-2023 raw GHCN warming trend of +0.075 °C decade⁻¹ is much larger than the T_{UHI} trend (+0.006 °C decade⁻¹), making the T_{UHI} trend only about 8% of the observed T_{AVG} trend. Thus, rural U.S. stations have a small UHI warming influence compared to the observed warming trend. The 6th order polynomial fit in Fig. 6a (and following figures) is added to better visualize the observed warming up to the 1930s, cooling to the 1970s, and warming thereafter, structures which are not present in the UHI warming curves.

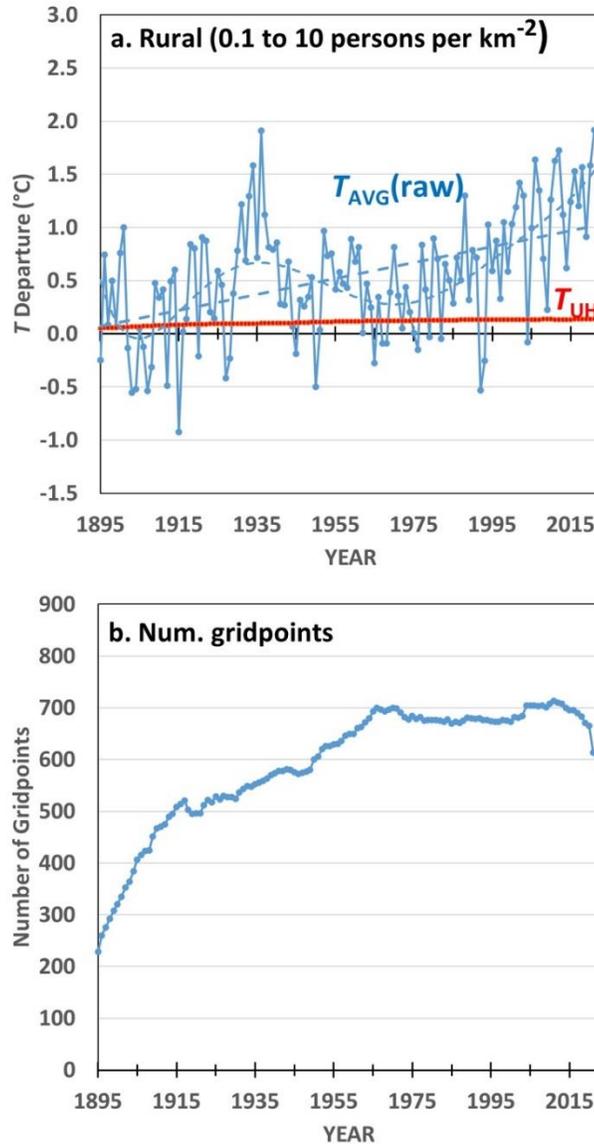


Fig. 6. CONUS area averages of cumulative year-on-year temperature departures from 1895 values in (a) summer (JJA) GHCN raw T_{AVG} and calculated T_{UHI} warming from dT_{AVG}/dPD calculations in Fig. 4 applied to station population density histories (Eq. 1), and (b) the number of 1 deg grid cells having station data for the 0.1 to 10 persons km⁻² (rural) population density category. The T_{AVG} (raw) curve in (a) has been vertically offset so that its linear trend line meets the T_{UHI} value in 1895. Also shown in (a) is a 6th order polynomial fit to the T_{AVG} data.

We conclude that the most rural stations have little spurious warming from UHI effects, at least as quantified using population density as a proxy for urbanization.

For the next larger population density category (peri-rural, 10 to 100 persons km⁻²), Fig. 7 shows a substantially larger proportion (42%) of the raw T_{AVG} trend (+0.049 °C decade⁻¹) being accounted for by the UHI trend (+0.020 °C decade⁻¹). Note, however, that the observed warming since the 1970s diverges from the UHI warming curve, suggesting that the observed warming trend since the 1970s has relatively little influence from the UHI effect. This also means that warming from 1895 to the 1970s can be completely explained by the UHI effect, as seen in Fig. 7a.

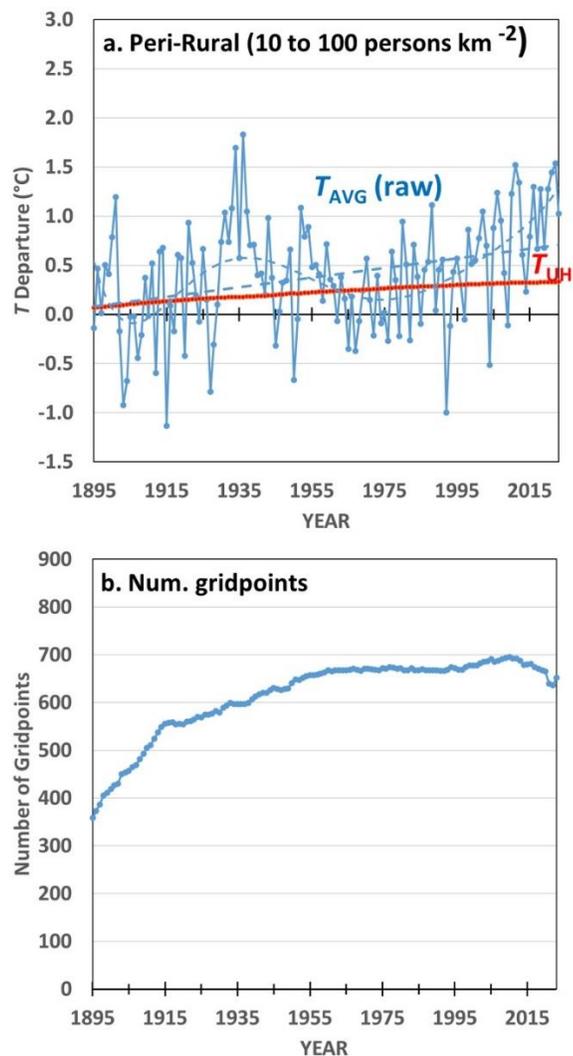


Fig. 7. As in Fig. 6, but for the 10 to 100 persons km⁻² (peri-rural) population density category.

For the 100-1,000 persons km⁻² (suburban) category, Fig. 8 shows a larger raw T_{AVG} warming trend (+0.065 °C decade⁻¹), with a still larger proportion (64%) being accounted for by the UHI trend (+0.041 °C decade⁻¹). Thus, the UHI effect is a substantial fraction of the centennial-scale warming observed in the raw temperature data. Again, however, the observed warming after the 1970s (as evidenced by the polynomial fit) is much larger than the UHI warming curve, suggesting little UHI influence on the raw temperature trend after the 1970s. In contrast, the observed warming from 1895 to the 1970s falls below the T_{UHI} curve, indicating comparable magnitudes of net warming in the first two-thirds of the temperature record.

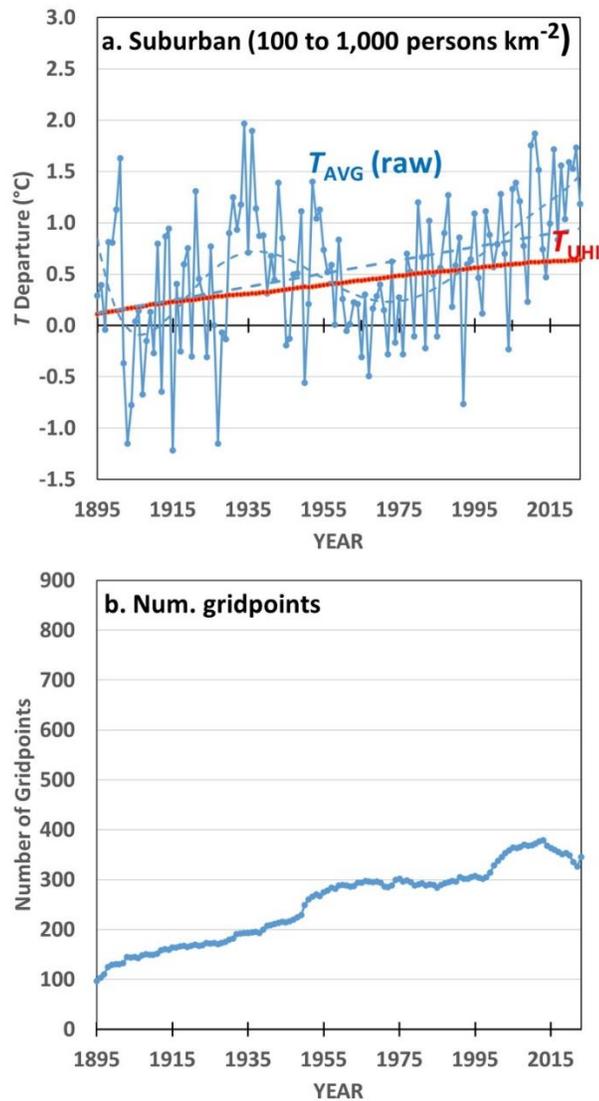


Fig. 8. As in Fig. 6, but for the 100 to 1,000 persons km⁻² (suburban) population density category.

In the fourth and final *PD* category, the urban category ($>1,000$ persons km^{-2} , Fig. 9) has relatively few stations, and so represents only a small portion of the U.S. But for those urban stations the T_{UHI} warming trend ($+0.078$ $^{\circ}\text{C}$ decade $^{-1}$) is a large fraction (67%) of the observed raw warming trend ($+0.117$ $^{\circ}\text{C}$ decade $^{-1}$). The net T_{UHI} warming amounts to about 1.1 deg. C over the period 1895-2023, and the T_{UHI} curve does not depart substantially from the polynomial fit to the raw temperature data until after 2000. This suggests that UHI effects should be considered and included when discussing unusually warm summer temperatures in urban areas.

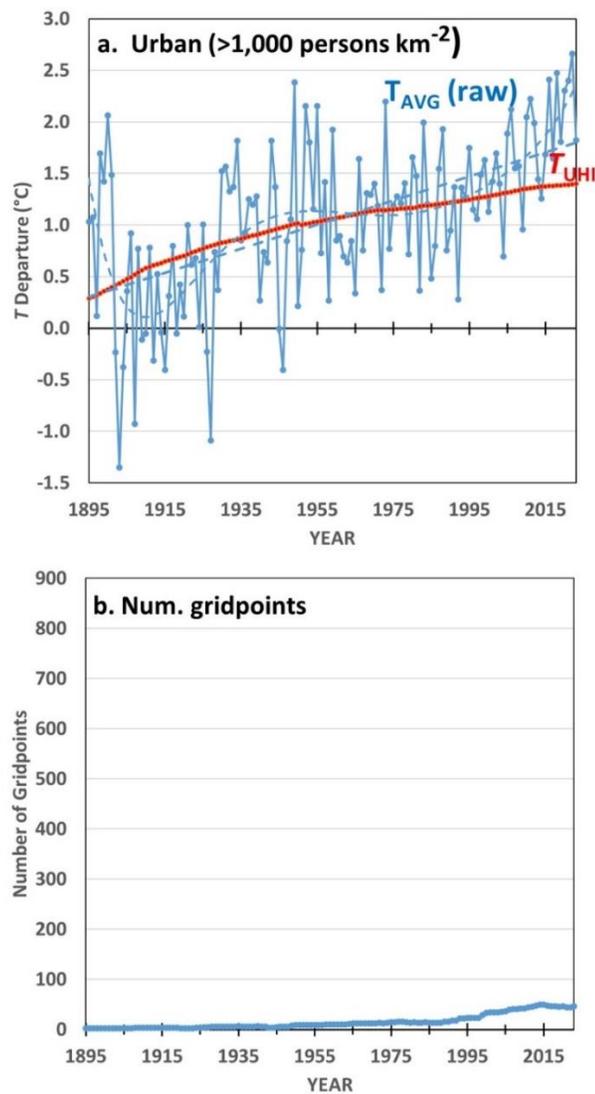


Fig. 9. As in Fig. 6, but for the $>1,000$ persons km^{-2} population density category.

Finally, we can use all the available stations (Fig. 10) to provide some idea of how much of the summertime CONUS warming between 1895 and 2023 in the raw GHCN data might be explained by the UHI effect.

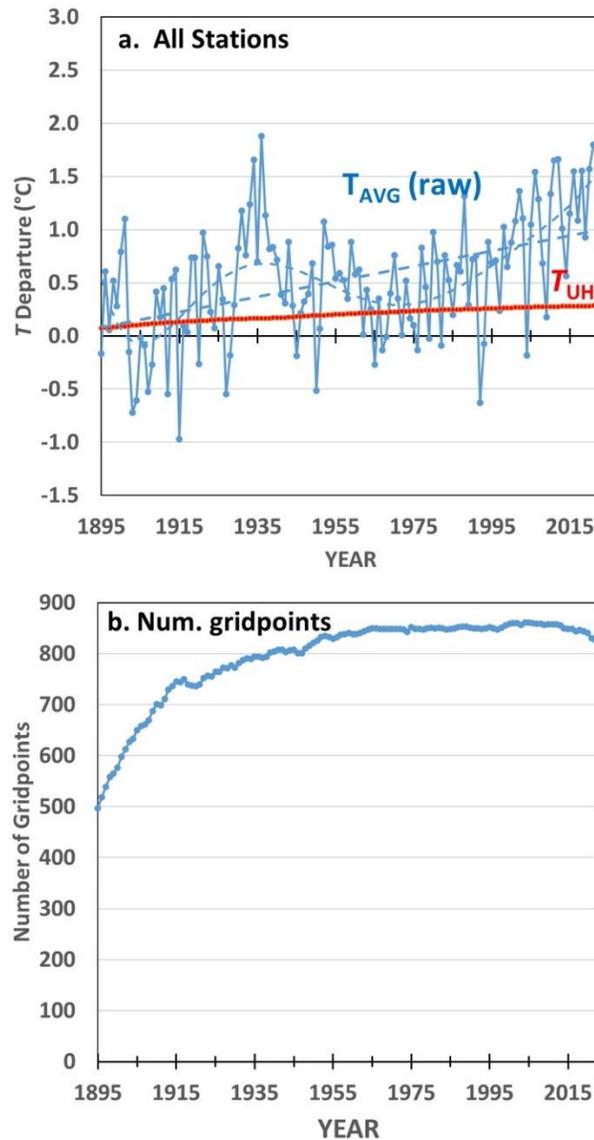


Fig. 10. As in Fig. 6, but for all stations. Other than an offset for display purposes, the T_{AVG} data in (a) are the same as in Fig. 2. The number of gridpoints is not the sum of those in Figs. 6 through 9 because gridpoints often have more than one population class of station represented.

Since the CONUS station mix is dominated by rural and peri-rural stations, the strong UHI effects seen in the two highest population density categories (Figs. 8 and 9) are reduced in the all-station average, leading to the CONUS-average T_{UHI} trend of $+0.016 \text{ }^\circ\text{C decade}^{-1}$, which is 22% of the observed trend ($+0.072 \text{ }^\circ\text{C decade}^{-1}$) in the raw data. These results can be compared to the Hausfather et al. (2013) analysis of the homogenized temperature data during 1895-2010 which found that “urbanization accounts for 14–21% of the rise in unadjusted minimum temperatures since 1895 and 6–9% since 1960”.

4. Summary and Discussion

A novel method is presented for calculating the multi-station average urban heat island effect on surface air temperatures based upon raw (non-homogenized) GHCN station T_{AVG} data and population density data at $\sim 10 \times 10 \text{ km}$ resolution using closely spaced station pairs. The sensitivity of monthly station average air temperature to differences in CONUS GHCN station population density (dT_{AVG}/dPD) for stations between 1880 and 2020 for the months of June, July, and August is estimated using regression between station-pair temperature and population differences in six separate classes of 2-station average population density. Every one of the six regressions in all 22 historical periods analyzed produce positive dT_{AVG}/dPD relationships (Table 2), resulting in a significant UHI signal in the GHCN data when averaged over all U.S. stations. As reported by Oke (1973) and others, we find the relationship between UHI and population to be nonlinear, with the greatest temperature sensitivity to population change at the lowest population densities (Fig. 1). This suggests that small absolute increases in population density at rural locations have potentially non-trivial impacts on temperature.

For a hypothetical time-invariant population density, we find that spatial UHI effects were somewhat stronger in the earlier decades and gradually weaken over time, especially at the more urban locations. We speculate this is due to better siting of thermometers over time which reduces microclimate effects on the thermometer measurements, an effect that is not captured by our use of low-resolution ($\sim 10 \times 10 \text{ km}$) population density data, although other explanations are also possible.

When GHCN stations are grouped into four general categories of station population density (0.1 to 10, 10 to 100, 100 to 1,000, and over 1,000 persons km^{-2}) we find average population increases since 1880 in all four categories. When the calculated dT_{AVG}/dPD sensitivity

relationships are applied to the station population changes over time using Eqs. (1) and (2), the cumulative T_{UHI} warming from 1880 to 2023 ranges from 0.14 °C for the most rural category of station (0.1 to 10 persons km⁻²) to 1.40 °C for the most densely populated urban locations (>1,000 persons km⁻²), with an all-station area average net UHI warming of 0.29 °C. The resulting T_{UHI} linear trends (1895-2023) range from 8% of the T_{AVG} trends for the most rural category (+0.006 versus +0.075 °C decade⁻¹) to 67% of the T_{AVG} temperature trend for the urban class (+0.078 versus +0.117 °C decade⁻¹). Averaged across all stations, the T_{UHI} linear trend of +0.016 °C decade⁻¹ is 22% of the observed raw T_{AVG} trend of +0.072 °C decade⁻¹. These results can be considered similar to those of Hausfather et al. (2013) who found that 14-21% of the warming in T_{MIN} data during 1895-2010 was due to the UHI effect, with a lesser fraction during 1960-2010. Since the UHI effect is known to be much stronger in daily minimum (T_{MIN}) than maximum (T_{MAX}) temperatures, our results for T_{AVG} will underestimate the UHI effects on T_{MIN} .

Taken together, our results suggest the need for more thorough efforts to account for urbanization effects over time when calculating land-based temperature statistics from GHCN data, especially when dealing with stations having population density over 100 persons km⁻² (suburban and urban). This does not necessarily mean that urbanization effects should be removed from station data since they represent the actual temperatures experienced by a large fraction of the U.S. population which resides in urban areas. Instead, at a minimum, the reporting and significance of suburban and urban temperature statistics (e.g. record high temperatures) should take into account urbanization effects, in addition to large-scale climate change, when those statistics are discussed. While our method for estimation of station-average UHI warming is useful in its own right, its relationship to commonly reported temperature trends depends upon the question of whether the adjusted (homogenized) GHCN data already have spurious UHI warming effects on temperature trends mostly accounted for. For the period 1895-2023 the raw GHCN trend (+0.072 °C decade⁻¹) is very close to the adjusted (homogenized) trend (+0.073 °C decade⁻¹), from which one might infer little UHI adjustment to the homogenized GHCN dataset. But UHI effects on trends in the homogenized data can be offset by other adjustments, for example the time of observation (TOBs, Vose *et al.*, 2003) adjustment, and station instrumentation and location changes. It is beyond the scope of this study to evaluate the relative sizes of these various adjustments since the homogenization algorithm makes no distinction between sources of discontinuities in the temperature data. We instead offer the current T_{UHI}

estimates as a way to isolate and quantify average UHI effects, and we obtain magnitudes that exceed 50% of reported trends for non-rural population densities. While our estimates of UHI warming are similar to those of Hausfather *et al.* (2013) in terms of temperature trends (which are presumably largely removed in the homogenized data), our estimate of UHI impacts on rural temperature trends is a new feature which results from the strong nonlinearity of the relationship between temperature and population density. Thus, the common assumption that rural temperature trends are unaffected by UHI is not strictly true.

Finally, it should be noted that our use of population density as a metric of urbanization is due to the ready availability of global gridded data, but it is also possible for increased UHI effects even after urban population densities have stabilized, due to expansion of built-up areas, paved surfaces, and energy use (Böhm, 1998). Thus, it is possible our estimates of T_{UHI} warming based upon population density alone will underestimate the total UHI effect.

Acknowledgments.

We thank three anonymous reviewers for many helpful suggestions. This research was funded through U.S. Department of Energy contract DE-SC0019296 and the Alabama Office of the State Climatologist.

Data Availability Statement.

GHCN Version 4 monthly data are available at <https://www.ncei.noaa.gov/pub/data/ghcn/v4/>. HYDE3.2 population density data are available at <https://public.yoda.uu.nl/geo/UU01/MO2FF3.html>. Derived datasets provided upon reasonable request.

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. *Int. J. Climatol.* **23**, 1-26, <https://doi.org/10.1002/joc.859>
- Balchin, W. G. V. and N. Pye, 1947: A micro-climatological investigation of Bath and the surrounding district. *Quart. J. Roy. Meteor. Soc.* **73**, 297–323.
- Böhm, R.. 1998: Urban bias in temperature time series – a case study for the city of Vienna, Austria. *Climatic Change* **38**, 113–128, <https://doi.org/10.1023/A:1005338514333>
- Chen, K., J. Boomsma, and H. A. Holmes, 2023: A multiscale analysis of heatwaves and urban heat islands in the western U.S. during the summer of 2021. *Sci. Rep.* **13**, 9570, <https://doi.org/10.1038/s41598-023-35621-7>
- Deming, W. E., 1943: *Statistical adjustment of data*. Wiley, NY (Dover Publications edition, 1985). ISBN 0-486-64685-8.
- Estoque, R.C., M. Ooba, X.T. Seposo, T. Togawa, Y. Hijioka, K. Takahashi, and S. Nakamura, 2020: Heat health risk assessment in Philippine cities using remotely sensed data and social-ecological indicators. *Nat. Commun.* **11**, 1581, <https://doi.org/10.1038/s41467-020-15218-8>
- Freire S., K. MacManus, M. Pesaresi, E. Doxsey-Whitfield, and J. Mills, 2016: Development of new open and free multi-temporal global population grids at 250 m resolution. Geospatial Data in a Changing World; Association of Geographic Information Laboratories in Europe (AGILE), AGILE 2016.
- Frost, C., and S. G. Thompson, 2000: Correcting for regression dilution bias: comparison of methods for a single predictor variable, *J. Royal Stat. Soc. Ser. A* **163**, 173-189, <https://doi.org/10.1111/1467-985X.00164>
- Fuller, W. A., 1987: *Measurement Error Models*. New York: Wiley. ISBN 9780470317334

- Gallo, K. P., T. W. Owen, D. R. Easterling, and P. F. Jamason, 1999: Temperature trends of the U.S. Historical Climatology Network based on satellite-designated land use/land cover. *J. Clim.* **12**, 1344–1348.
- Gallo, K.P, and T. W. Owen, 2002: A sampling strategy for satellite sensor-based assessments of the urban heat-island bias. *Int. J. Rem. Sens.* **23**, 1935-1939, <https://doi.org/10.1080/01431160110097259>
- Goldewijk, K., A. Beusen, J. Doelman, and E. Stehfest, 2017: Anthropogenic land use estimates for the Holocene; HYDE 3.2, *Earth Sys.Sci. Data* **9**, 927-953, <https://doi.org/10.5194/essd-9-927-2017>
- Hamdi, R., and H. Van deVyver, 2011: Estimating urban heat island effects on near-surface air temperature records of Uccle (Brussels, Belgium): an observational and modeling study. *Adv. Sci. Res.* **6**, 27-34, <https://doi.org/10.5194/asr-6-27-2011>
- Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change, *Rev. Geophys.* **48**, RG4004, 29 pp.
- Hausfather, Z., M.J. Menne, C.N. Williams, T. Masters, R. Broberg, and D. Jones, 2013: Quantifying the effect of urbanization on U.S. Historical Climatology Network temperature records. *J. Geophys. Res. – Atmos.* **118**, 481-494, <https://doi.org/10.1029/2012JD018509>
- Hsu, A., G. Sheriff, T. Chakraborty, and D. Manya, 2021: Disproportionate exposure to urban heat island intensity across major US cities. *Nat Commun* **12**, 2721, <https://doi.org/10.1038/s41467-021-22799-5>
- Hutcheon, J. A., A. Chiolerio, and J. A. Hanley, 2010: Random measurement error and regression dilution bias. *Brit. Med. J.* **340**, c2289–1406, <https://doi.org/10.1136/bmj.c2289>
- Jones, P. D., D. H. Lister, and Q. Li, 2008: Urbanization effects in large-scale temperature records, with an emphasis on China, *J. Geophys. Res.* **113**, D16122, <https://doi.org/10.1029/2008JD009916>
- Karl, T. R., H. F. Diaz, and G. Kukla, 1988: Urbanization: Its detection and effect in the United States climate record. *J. Clim.* **1**, 1099–1123, [https://doi.org/10.1175/1520-0442\(1988\)001<1099:UIDAEI>2.0.CO;2](https://doi.org/10.1175/1520-0442(1988)001<1099:UIDAEI>2.0.CO;2)

- Katata, G., R. Connolly, and P. O'Neill, 2023: Evidence of urban blending in homogenized temperature records in Japan and in the United States: Implications for the reliability of global land surface air temperature data. *J. Appl. Meteor. Climatol.*, **62**, 1095–1114, <https://doi.org/10.1175/JAMC-D-22-0122.1>
- Kukla, G., J. Gavin, and T.R. Karl, 1986: Urban warming. *J. Clim. Appl. Meteor.* **25**, 1265–1270.
- McKittrick, R., 2023: On the choice of TLS versus OLS in climate signal detection regression. *Clim. Dyn.* **60**, 359–374, <https://doi.org/10.1007/s00382-022-06315-z>
- Menne, M. J., and C. N. Williams, 2009: Homogenization of temperature series via pairwise comparisons, *J. Clim.*, **22**, 1700–1717, <https://doi.org/10.1175/2008JCLI2263.1>
- Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston, 2012: An overview of the Global Historical Climatology Network-Daily database. *J. Atmos. Oceanic Technol.* **29**, 897–910, <https://doi.org/10.1175/JTECH-D-11-00103.1>
- Menne, M. J., C. N. Williams, B. E. Gleason, J. J. Rennie, and J. H. Lawrimore, 2018: The Global Historical Climatology Network Monthly temperature dataset, Version 4. *J. Clim.* **31**, 9835–9854, <https://doi.org/10.1175/JCLI-D-18-0094.1>
- Mills, G., 2008: Luke Howard and the climate of London. *Weather - Royal Met. Soc.* **63**, 153–157, <https://doi.org/10.1002/wea.195>
- Mitchell, J. M., 1953: On the causes of instrumentally observed secular temperature trends. *J. Atmos. Sci.* **10**, 244–261.
- Oke, T.R., 1973: City size and the urban heat island, *Atmos. Env.* **7**, 769–779.
- Oke, T. R., 1982: The energetic basis of the urban heat island. *Quart. J. Royal Met. Soc.* **108**, 1–24.
- O'Neill, P., and Coauthors, 2022: Evaluation of the homogenization adjustments applied to European temperature records in the Global Historical Climatology Network dataset. *Atmos.* **13**, 285, <https://doi.org/10.3390/atmos13020285>
- Parker, D. E., 2006: A demonstration that large-scale warming is not urban. *J. Clim.* **19**, 2882–2895, <https://doi.org/10.1175/JCLI3730.1>

- Parker, D.E., 2010: Urban heat island effects on estimates of observed climate change. *Wiley Interdisciplinary Reviews - Climate Change*, **1**, 123-133, <https://doi.org/10.1002/wcc.21>
- Peterson, T.C., and Coauthors, 1999: Global rural temperature trends. *Geophys. Res. Lett.* **26**, 329-332, <https://doi.org/10.1029/1998GL900322>
- Peterson, T. C., 2003: Assessment of urban versus rural in situ surface temperatures in the contiguous United States: no difference found. *J. Clim.*, **16**, 2941–2959, [https://doi.org/10.1175/1520-0442\(2003\)016<2941:AOUVRI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<2941:AOUVRI>2.0.CO;2)
- Peterson, T. C., and T. W. Owen, 2005: Urban heat island assessment: Metadata are important. *J. Clim.* **18**, 2637–2646, <https://doi.org/10.1175/JCLI3431.1>
- Scafetta, N., 2021: Detection of non-climatic biases in land surface temperature records by comparing climatic data and their model simulations. *Clim. Dyn.* **56**, 2959–2982, <https://doi.org/10.1007/s00382-021-05626-x>
- Soon, W., R. Connolly, M. Connolly, P. O’Neill, Z. Jingyun, Q. Ge, Z. Hao, and H. Yan, 2018: Comparing the current and early 20th century warm periods in China, *Earth-Sci. Rev.* **185**, 80-101, <https://doi.org/10.1016/j.earscirev.2018.05.013>
- Spearman, C., 1904: The proof and measurement of association between two things. *Am. J. Psych.*, **15**, 72–101.
- Stewart, I.D. and T. R. Oke, 2012: Local climate zones for urban temperature studies. *Bull. Amer. Meteor. Soc.* **93**, 1879-1900, <https://doi.org/10.1175/BAMS-D-11-00019.1>
- Summers, P. W., 1964: An urban ventilation model applied to Montreal. Ph.D. Thesis, McGill University.
- Tong, S, J. Prior, G. McGregor, X. Shi, and P. Kinney, 2021: Urban heat: an increasing threat to global health *Brit. Med. J.*, **375**:n2467, <https://doi.org/10.1136/bmj.n2467>
- Vose, R., C. Williams, T. Peterson, T.R. Karl, and D. Easterling, 2003: An evaluation of the time of observation bias adjustment in the US Historical Climatology Network. *Geophys. Res. Lett.* **30**, <https://doi.org/10.1029/2003GL018111>
- Vose, R. S., S. Applequist, M. Squires, I. Durre, M. J. Menne, C. N. Williams, C.

Fenimore, K. Gleason, and D. Arndt, 2014: NOAA Monthly U.S. Climate Divisional Database (NClimDiv). NOAA National Climatic Data Center.

<https://doi.org://10.7289/V5M32STR>

Wickham, C., and Coauthors, 2013: Influence of urban heating on the global temperature land average using rural sites identified from MODIS classifications. *Geoinfor. Geostat. An Overview* **1**, <https://doi.org/10.4172/2327-4581.1000104>.

Zhang, K., C. Cao, H. Chu, L. Zhao, and X. Lee, 2023: Increased heat risk in wet climate induced by urban humid heat. *Nature*, **617**, 738–742, <https://doi.org/10.1038/s41586-023-05911-1>