ORIGINAL PAPER

Detecting impacts of surface development near weather stations since 1895 in the San Joaquin Valley of California

Diane Kim[1,2] · John R. Christy[1]

Abstract

Temperature readings observed at surface weather stations have been used for detecting changes in climate due to their long period of observations. The most common temperature metrics recorded are the daily maximum (TMax) and minimum (TMin) extremes. Unfortunately, influences besides background climate variations impact these measurements such as changes in (1) instruments, (2) location, (3) time of observation, and (4) the surrounding artifacts of human civilization (buildings, farms, streets, etc.) Quantifying (4) is difficult because the surrounding infrastructure, unique to each site, often changes slowly and variably and is thus resistant to general algorithms for adjustment. We explore a direct method of detecting this impact by comparing a single station that experienced significant development from 1895 to 2019, and especially since 1970, relative to several other stations with lesser degrees of such development (after adjustments for the (1) to (3) are applied). The target station is Fresno, California (metro population ~ 15,000 in 1900 and ~ 1 million in 2019) situated on the eastern side of the broad, flat San Joaquin Valley in which several other stations reside. A unique component of this study is the use of pentad (5-day averages) as the test metric. Results indicate that Fresno experienced + 0.4 °C decade^{-1} more nighttime warming (TMin) since 1970 than its neighbors—a time when population grew almost 300%. There was little difference seen in TMax trends between Fresno and non-Fresno stations since 1895 with TMax trends being near zero. A case is made for the use of TMax as the preferred climate metric relative to TMin for a variety of physical reasons. Additionally, temperatures measured at systematic times of the day (i.e., hourly) show promise as climate indicators as compared with TMax and especially TMin (and thus TAvg) due to several complicating factors involved with daily high and low measurements.

Keywords Temperature · Urbanization · Climate · Time series construction

1 Introduction

To address the widely publicized issue of calculating the magnitude of the response of the climate system to human-caused increases in greenhouse gases (GHGs), it is fundamentally necessary to utilize observations which describe what has happened to the climate system from periods before and during the period of rising GHGs. Surface temperature is de facto one of the key variables because observations are available back into the nineteenth century before the response of the climate system to extra GHGs would have

been significant. While possessing the key trait of providing a long-term record of a climate-response variable, it is one that is unfortunately confounded by other signals since it also responds to non-GHG effects, e.g., urbanization, and the vagaries due to changes in equipment, observing practices, exposure, time of observation, and location (Thorne et al. 2011 and citations therein). Not to be overlooked is a more fundamental confounding factor in that the essential character of the climate system is its nonlinear dynamical behavior, having the capability to generate long-term variations without the need for external forcing, such as GHGs, which can naturally cause excursions outside of those observed in our relatively short record of ~ 125 years.

Numerous studies have delved into the complex issues raised by the fact that surface temperature observations are beset by numerous inhomogeneity issues. Many of these problems are not well characterized, so that they render the construction of a century-scale perfectly "pristine" time

✉ John R. Christy
   christy@nsstc.uah.edu

1 The University of Alabama in Huntsville, Huntsville, AL, USA

2 Duke University, Durham, NC, USA

series essentially impossible though many homogenization methods have been attempted (e.g., Karl et al. 1986; Parker 1994; Christy et al. 2006; Pielke, Sr. et al. 2007; Menne and Williams 2009; Williams et al. 2012 and citations therein). Indeed, there is a vast body of literature in which many methods of data adjustment are published to deal with "extraneous biases" in these temperature datasets (e.g., McKitrick and Michaels 2007).

In this study we shall describe the unadjusted datasets we utilized and then the adjustments that were necessarily applied in an attempt to remove the specific inhomogeneities due to sudden events such as instrument and location changes. It is important to note that only limited documentation is available regarding the changes in instruments, exposures, and practices. Yet even when available, this information does not tell us what the impact such changes have on the temperature metric itself. As a result, we must use objective techniques to detect temperature shifts which arise from events which may or may not be documented. From this will be produced an adjusted time series for Fresno from which the estimated impact of urbanization will be calculated, which is the goal of this study. (Note we refer to our changes as "adjustments" rather than "corrections" because we are never certain of their complete accuracy.)

Of particular novelty in this project is the use of the pentad time scale as the temporal metric. By averaging the daily values into pentads (5 days), we reduce the noise of high-frequency weather variations as well as the random error associated with measurements such as these. Additionally, the annual cycle is more accurately determined as there are 73 representative points rather than the typical 12 monthly points. There were 9125 pentads in the 125 years covering 1895 to 2019. Another new approach is introduced with the use of hourly data, available back to 1893 in Fresno.

2 Data sources

The National Center for Environmental Information (NCEI) houses the largest archive of climate data in the world. Much of the information originally recorded on paper forms has been manually keyed into computer-readable digital files. Such archives have provided access to data for studies that in the last century would have taken enormous amounts of time just to put the data in useable form. After examining time series of about 50 stations, we selected in addition to Fresno, fifteen nearby stations, Angiola, Clovis (near), Friant, Hanford, Le Grand, Lemon Cove, Lindsay, Madera, Merced, Newman, Porterville, Reedley, Tulare, Visalia, and Wasco (Fig. 1). These fifteen stations were selected based on the availability of a relatively long period of data, their proximity to Fresno, and that they also reside in the floor of the San Joaquin Valley. The metric we employ will be the daily extreme temperatures commonly known as the daily high (TMax) and daily low (TMin).

Of particular interest is that three stations, Clovis-near (1916–1947), Reedley (1895–1920), and Tulare (1895–1906) were not available as computer readable files. The daily observations for these three stations were manually keyed-in from images of the original documents

Fig. 1 Map of the central San Joaquin Valley of California with stations used in this study

archived at NCEI and provided important information for the early period—a total of 49,541 daily values otherwise not available from NCEI. Details of these stations are provided in the book "*Is it getting hotter in Fresno … or not?*" by the second author in which a different method of dataset construction was applied to a different set of stations than utilized in this investigation (Christy 2021). The data availability is shown in Fig. 2.

3 Method

In the methodology that follows, we introduce an assumption that the weather records in Fresno were measured and reported accurately. Fresno was an official Weather Bureau/National Weather Service station from the late 1800s, was provided with standard instrumentation, was staffed with federal observers, and recorded readings for the calendar day, i.e., midnight to midnight. We will consider Fresno the target station for our analysis. Though the temperatures were measured and recorded in degrees Fahrenheit, we shall convert all to degrees Celsius.

The unadjusted data present many challenges for studies such as this. Our first task was to account for time-of-observation bias in the non-Fresno stations. Some observations were recorded in the morning at 0700 h, others in the afternoon around 1800 h and Fresno (with other stations when they were moved to airports) at midnight. However, documentation regarding the time of observation is often missing, so we used a statistical method to determine at what time of day the observation was likely taken and made adjustments accordingly. In this process we are attempting to convert the non-midnight station values to values as would have been observed had the station been a midnight station—in this way we may compare stations having reduced the uncertainty due to this feature.

The main issue here is that a station reading TMax and TMin for the past 24 h at 0700 h will, in the vast majority of cases, be recording TMax that occurred the previous day. Observations taken at 1800 h will likely capture the same TMax and TMin as the midnight station since the values of TMax and TMin generally occur between 0000 h of the day in question and current late afternoon. Though this is not always the case as TMax and TMin may occur at any hour of the day, the dominate time for TMax is mid-afternoon
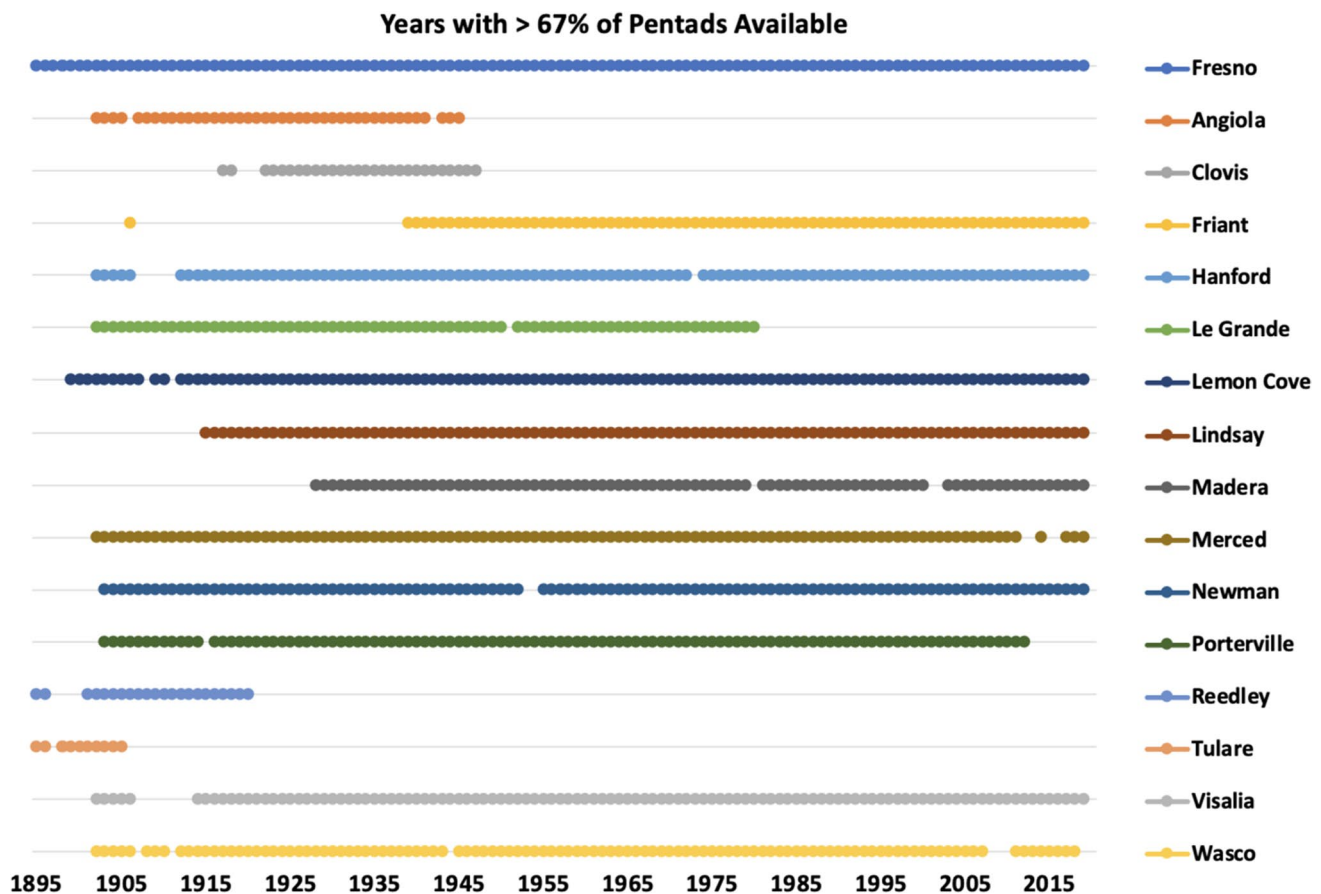

Fig. 2 Chart of data availability by year. Symbol appears if > 67% of pentads (at least 49) contained all 5 days of data

and for TMin, near sunrise. Thus, as we adjust the readings, we realize they will not produce a correct calendar day (i.e., midnight observation) representation in every case.

To address this issue for the 0700 h stations, we compared the daily TMax values of each station to the target station, Fresno, whose observations were consistently recorded at midnight. Using a 15-day window, we generated the correlations between daily TMax values of Fresno against each of the fifteen stations. For each station, we then shifted the window back by 1 day and again calculated the correlation between Fresno's TMax values and the station's shifted values. [Note: we are using daily values, not pentads.] If the correlations between Fresno and the shifted window were higher than the correlations between Fresno and the non-shifted window for a set period of time, then the observations in that station must be shifted back in order to account for the time of observation bias. Thus, for each month, we shifted TMax values back by 1 day if their correlations with Fresno were higher since this most likely indicated the station took observations around 0700 h.

Issues related to observations at 1800 h will be addressed through the removal of breakpoint biases. For example, an observer recording at 1800 h will read the highest temperature measured in the previous 24 h, i.e., from 1800 h the day before to 1800 h today. If the preceding day was hot, i.e., the value at 1800 h was hot, followed by a cold front later in the evening, the highest temperature attained after the front's passage on the following calendar day would be much cooler. Yet, the observer would read a TMax value that had been attained at 1800 h the day before when the thermometer was reset rather than the TMax that would have occurred since midnight as recorded in the Fresno station. This "double counting" of hot days for 1800 h stations creates a bias in TMax values and the intent will be that it be removed during the breakpoint adjustment procedure.

After these adjustments, we then performed an outlier-removal sweep to clear the dataset of highly unlikely temperature readings. This was done by calculating the mean and standard deviation of the set of temperature values for each day of the calendar year, converting the actual temperature anomaly (specific value minus mean for that calendar day) to a z-score (anomaly divided by standard deviation). We then checked the value of the z-score of the station against the z-score of Fresno for the same day. This was performed separately for TMax and TMin. If the z-score was greater than 2.5 and the difference in z-scores was greater than 2, the non-Fresno station data value was set to missing. Note that we are removing anomalies which not only have a z-score of 2.5 or more, but also whose z-score *difference* with Fresno exceed the value of 2. This check alerted us, for example, to several erroneous values that appeared to have been keyed-in without the tens digit for TMin values. Since single-digit values (degrees Fahrenheit) have never been observed in Fresno, and the difference versus Fresno was typically on the order of 20 °F (11 °C) or more, these were obvious candidates for elimination.

To make the data more manageable and useful to our purpose, we generated pentad (5-day) mean values as our fundamental test metric. We thus had 73 pentads per year with February 29 in Leap Years included as a sixth observation in the 12th pentad. Difference time series for each station relative to Fresno was then calculated (station minus Fresno) to quantify the magnitude of the differences and their changes over time. Since the raw difference series are not an accurate measure of true climate-induced changes (there are still inhomogeneities to remove), several more statistical adjustments were necessary before seeking patterns in the data.

Due to each station's unique microclimate, adjustments for location-induced biases were needed. To this end we generated the mean annual cycle of each time series, using a Fourier series to create a smooth approximation of each curve, and subtracted from the processed time series their corresponding annual cycles to allow for a more direct comparison of the stations. In Fig. 3 we show the actual mean annual cycle of TMax and TMin of daily values for Fresno as well as the Fourier approximation which was used to generate the time series of anomalies. In Figs. 4a and 4b are the Fourier-smoothed mean annual cycle differences between the fifteen comparison stations and Fresno.

The results in Fig. 4a alert us of some regional variations in temperature patterns. Fresno TMin values are always warmer than the other stations, likely for three reasons. Firstly, until 1939, Fresno's official station was mounted on the roofs of various downtown buildings at elevations from 19 to 29 m above the ground since 1890. As is well known in meteorology, nighttime temperatures tend to be coldest at ground level as inversions often form due to rapid radiational cooling of the surface after sunset which causes the air near the ground to also cool. This produces a temperature profile that is cooler at the surface than at elevations several meters above (Walters et al. 2007). For locations in the broad San Joaquin Valley, nighttime inversions are present almost every night. Secondly, many of the stations are in lower-lying parts of the valley and subject to cold-air drainage more so than Fresno which sits on a slight ridge between the San Joaquin and Kings Rivers. The third factor is likely due to Fresno's rapid population growth after World War II and the influence of built-up infrastructure which was able to retain heat as well as inducing vertical mixing, again keeping TMin values warmer than they would be in a natural setting (Karl et al. 1993; Karl et al. 1998; Christy et al. 2006.). An examination of Porterville, for example, which is the coolest station relative to Fresno, suggests all three of these effects are operating; (1) cooler TMin due to Porterville's near-ground-level observations, (2) less urbanization, and (3) Porterville lies near the river bottom of the Tule River

Fig. 3 **a** Mean annual cycle of daily TMin for Fresno (unadjusted) both unsmoothed and approximated by Fourier analysis which is used for anomaly calculation. **b** Mean annual cycle of daily TMax for Fresno (unadjusted) both unsmoothed and approximated by Fourier analysis which is used for anomaly calculation
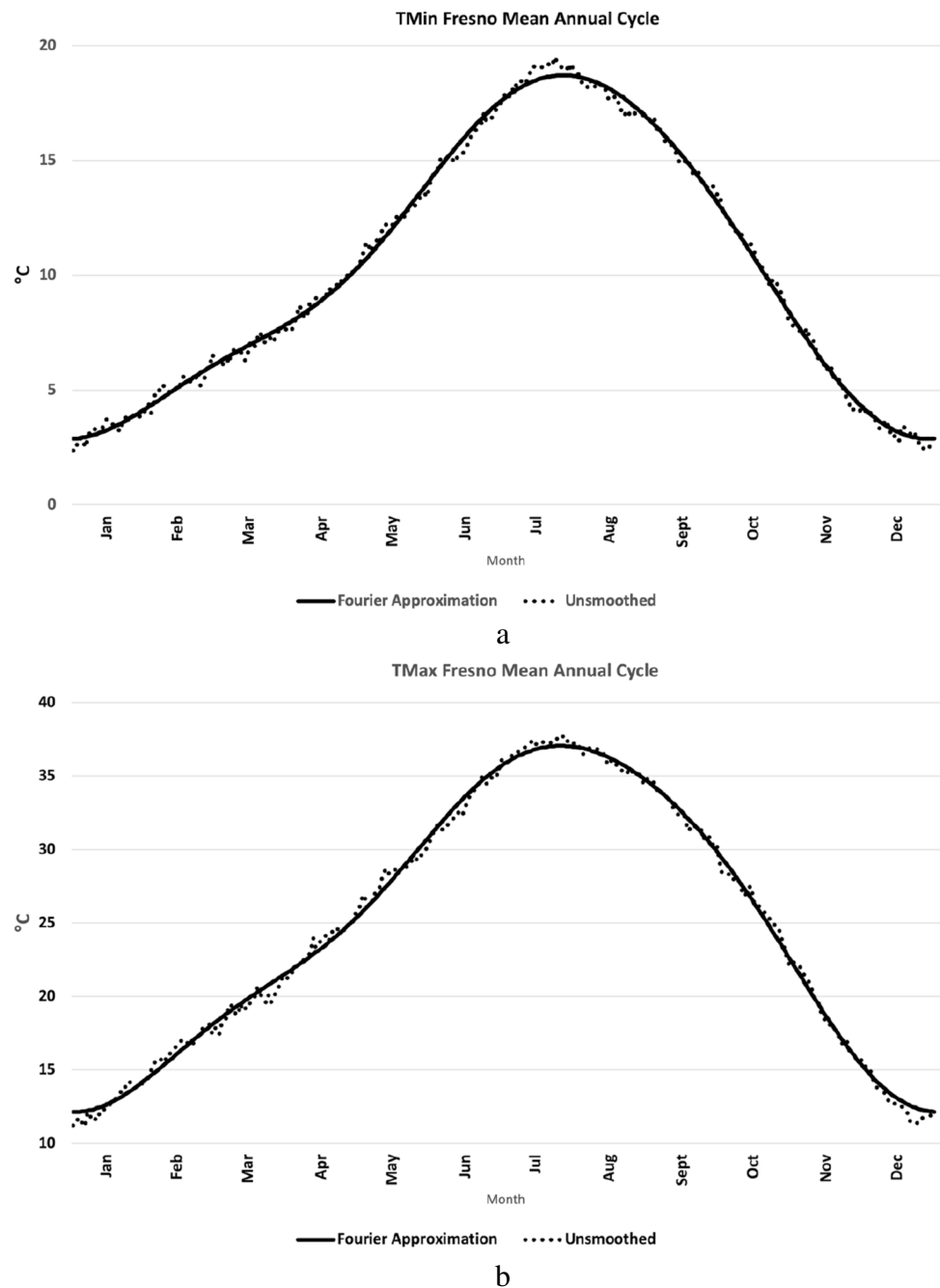

system affected by local cold-air drainage whereas Fresno is not.

With regard to TMax (Fig. 4b), Fresno tends to be slightly cooler on average than the other stations. The larger variations, such as Tulare, indicate the thermometer there was likely exposed to sunlight in a less than optimal siting situation, in which the cold season allowed greater direct exposure (i.e., a thermometer under an eve that allowed greater exposure when the sun was lower in the sky). In general, the assumption is that these variations are systematic and thus removed with the usage of anomalies as the test metric.

Adjusting for the differences in rooftop and ground level observations, which for Fresno occurred in July 1939, requires specific information. Griffith and McKee (2000) noted that the temperatures measured at rooftop and ground level vary from location to location so that each situation must be uniquely assessed and appropriately adjusted. This was possible for Fresno by using the breakpoint methodology with the several comparator stations.

Fig. 4 a Difference between mean annual cycle of Fresno and the fifteen comparison stations (station minus Fresno) based on the 73 pentad values. **b** As in Fig. 3a but for TMax
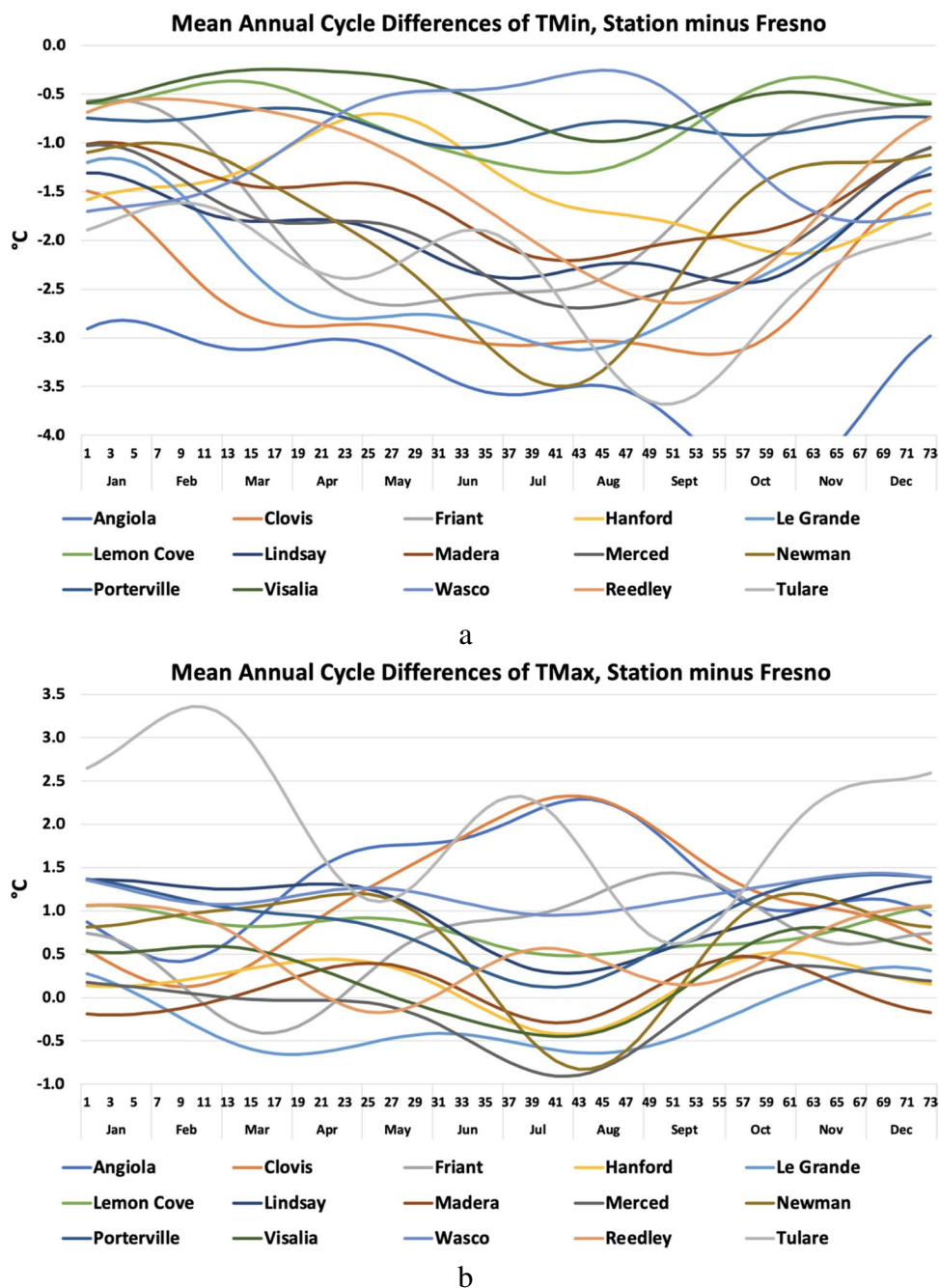
4 Non-climatic shifts in station observations

The goal here is to remove the impact of various incidents causing sudden, non-climatic shifts in the data (i.e., station relocation, change in instrument, moving station from top of building to ground level, etc.). As many of these shifts are undocumented, we chose to utilize a statistical technique to detect breakpoints and adjust each time series based on an intercomparison with other stations. The goal of this is to improve homogeneity of the data, thereby allowing climate-induced trends to emerge. The procedure we used relied on

this test statistic based on the average of the differences (μ) between stations:

$$\tau_k = \Delta\left\{\left[\mu_k^-(\Delta) - \mu_k(2\Delta)\right]^2 + \left[\mu_k^+(\Delta) - \mu_k(2\Delta)\right]^2\right\}/\sigma_k(2\Delta)$$

Where Δ is the half-window time width of the interval being examined, μ_k^- (μ_k^+) is the mean value of the temperature difference between stations in the first (last) half-window, μ_k is the mean of the complete 2Δ window, and σ_k is the standard deviation of the differences in the 2Δ window. If the test metric τ_k exceeds a certain significance threshold H, we

identified it as a breakpoint and shift all prior values by the value $\mu_k^+ - \mu_k^-$. The five values of H used were 200, 100, 75, 60, 50, and 35, but the key results of this study are based on the average of the H values of 75 and 60 together (see Haimberger 2007, Christy et al. 2009 for complete discussion of technique, and Christy and McNider 2016 for a discussion of tests on synthetic and reference datasets). Note that the test metric is the time series of the differences between two stations and H values of 100 and 75 correspond to z-scores of approximately 4.0 and 3.5, respectively.

We use an iterative process to determine the breakpoints for Fresno and the 15 non-Fresno stations referred to as "Valley." We first order the Valley stations by their data volume. There are 9125 pentads on the 1895 to 2019 period with Fresno reporting all 5 days in all pentads. Lemon Cove reported the most of the Valley stations with 8580 pentads on which all 5 days were available. Though Tulare reported the least with 756 pentads, its record began in 1895, hence its value in helping to test the earliest period.

Breakpoint values were determined and applied in this sequence. The first breakpoint for Fresno was determined by pair-wise comparisons with all of the Valley stations. Each of the fifteen comparisons produced a time series of H values based on the difference time series between Fresno and the individual Valley stations. The individual Valley time series of H were averaged from which the maximum value (i.e., most significant) was determined and used as the first breakpoint for Fresno. Fresno was then adjusted to remove this break identified by the average of the Valley stations. Next, this process was applied to Lemon Cove (the station with the most data of the remaining stations) but using the newly adjusted Fresno time series with the other fourteen unadjusted Valley stations. As before, the average time series of H was determined from all of the non-Lemon Cove stations to find the maximum H and then Lemon Cove was adjusted to remove this event.

The process was repeated for each station using the stations with longer records having their first breakpoint removed and the stations with fewer records not yet having their first breakpoint removed. At the end of this first sweep, all stations have their first breakpoint removed. The process is then repeated to discover the second, third, etc., breakpoints until no more values above the specific H threshold is found.

5 Results

Figures 5a and b display the TMin and TMax time series of the unadjusted annual anomalies for the Fresno and Valley stations. The graphs indicate general variations as well as likely points where discontinuities occur. Because the Valley time series uses 15 stations, the various discontinuities of the individual stations tend to average out, making this time series smoother and closer to what would be expected from an adjusted time series. In contrast, the single station of Fresno shows likely breakpoints in the 1940s and 1970s among other places for TMin as seen in the relative differences between the two time series. In Fig. 5b, early values of Valley TMax are considerably warmer than Fresno during a time when the Fresno station experienced few potential breaks and resided at a higher, cooler rooftop location. Otherwise, the two TMax time series demonstrate good agreement even in their unadjusted states. This is a common feature of surface temperature records in that TMin is the metric most affected by changes versus TMax (Christy et al 2006; McNider et al. 2012; Scafetta 2021).

The breakpoint adjustment procedure was applied to each individual station from $H = 200$ to $H = 35$. At $H = 35$ and 50, there were many breakpoints detected, some likely due to the random processes of interstation differences that naturally occur. With $H = 200$, there were no breakpoints detected for Fresno as the significance test allowed only extremely significant breakpoints to be accepted. A general rule is that surface stations of this type experience about one breakpoint per 7 to 15 years (Christy et al. 2006). The values of $H = 60$ and $H = 75$ generated time series with breakpoint incidences in line with the expected number; thus, we shall use the average of these two as that which eliminates (a) mostly true non-climatic breakpoints, and (b) the fewest breakpoints due to natural causes. Detailed discussion of the thresholds for these test metrics is found in Christy and McNider 2016.

Figure 6a shows the result for TMin as the adjusted time series of Fresno and the combination of the 15 comparison stations or Valley. Correlations between the two are 0.88 (0.77) for TMin (TMax). Comparing these to the unadjusted correlations of 0.76 (0.43) for TMin and Tmax, respectively, it is clear that the removal of breakpoints has improved the agreement, especially with TMax (Fig. 6b) due to the improvement in the early Valley values.

Our interest here is the detection of differences in the long-term temperature change between the two datasets. The 125-year trends for TMin for Fresno (Valley) are $+0.26 \pm 0.14$ ($+0.17 \pm 0.06$) °C decade^{-1}. More relevant are the trends from 1970 to the present, the period in which significant urban growth occurred around the Fresno station. Here the Fresno (Valley) trends are highly significantly different, $+0.63$ ($+0.22$) °C decade^{-1}, suggesting a value of the urban effect of approximately $+0.4$ °C decade^{-1} in the last 50 years—at least a lower bound as the other stations also certainly experienced some urban growth, as well.

The values for TMax indicate a very different outcome. For the 125-year period of record, Fresno (Valley) TMax trends are $+0.01 \pm 0.09$ (-0.05 ± 0.10) °C decade^{-1} and since 1970, $+0.27$ ($+0.31$) °C decade^{-1}. These differential trends are not statistically different from each other, and for

Unadjusted TMin Annual Values

a

Unadjusted TMax Annual Values

b

Fig. 5 **a** Time series of TMin annual anomalies of the unadjusted Fresno and Valley stations. **b** As in Fig. 4a but for TMax
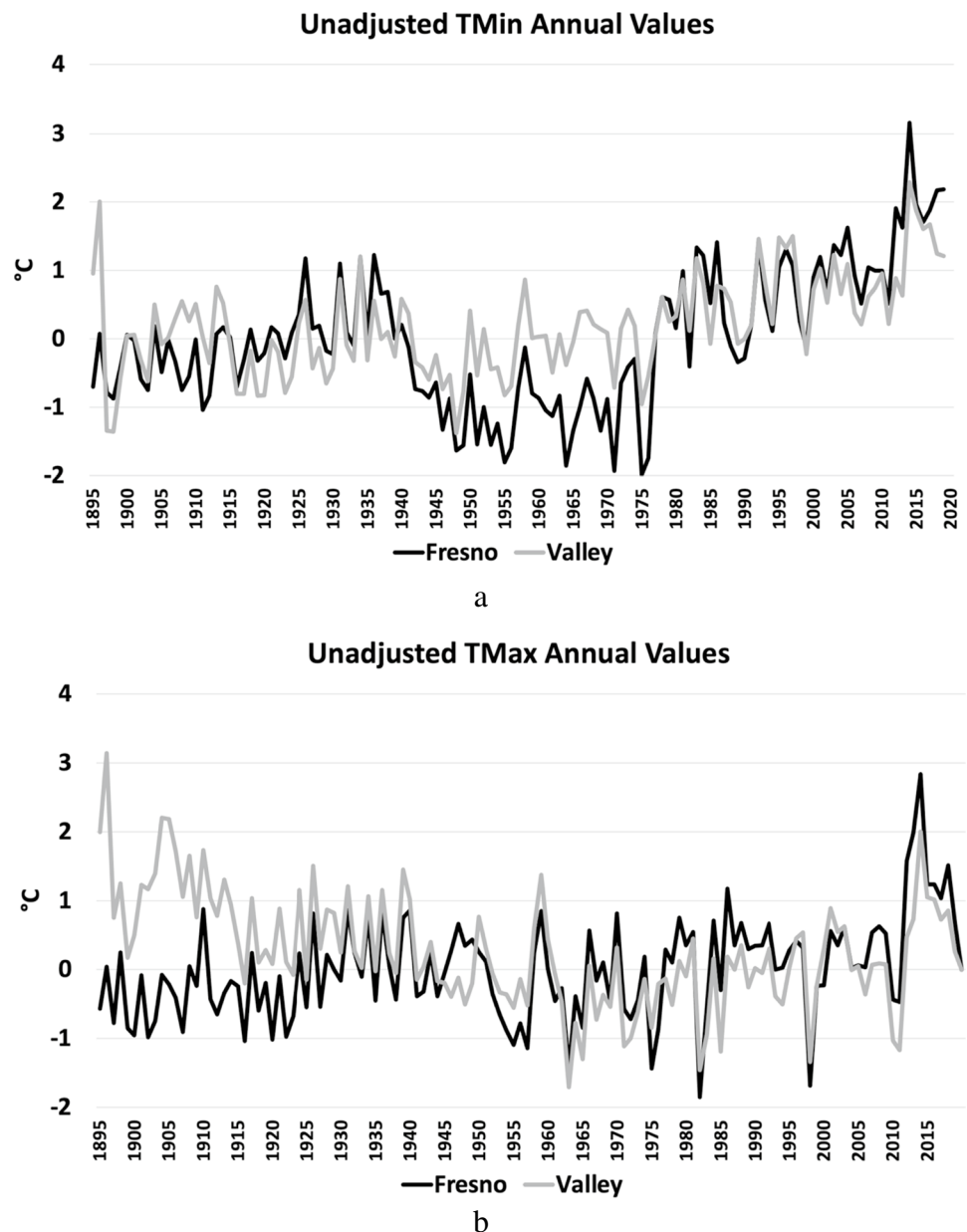
the full 125-year record are not statistically different from zero.

As a partially independent check, we obtained hourly temperatures for Fresno to compute 00UT and 12UT temperature time series representing 0400 and 1600 local standard time. We chose a single month, July, due to its relatively small interannual variability and the significant effort that was required to manually key-in the data.

The use of hourly data avoids the time-of-observation problem in which daily high and low temperatures can occur at any time in a 24-h period which can vary from station to station due to the fact that different stations recorded TMax and TMin for differing 24-h periods. As well, the more complicated mechanical issues (indexes that ride on top or

underneath the column of liquid that often malfunction) and digital aspects (variable time constants to determine high and low temperatures from observations taken at very small intervals) of determining TMax and TMin are completely avoided. Finally, the hourly reading was often read from a different thermometer than the TMax and TMin thermometers, as part of the psychometric calculations, so it is, again, a basically independent check.

We accessed these hourly values; Downtown Fresno (FNO, WBAN 53,125, 1895–1939), Chandler Field (FCH, WBAN 23,167, 1933–1949), and the Fresno-Yosemite International Airport (FAT, WBAN 93,193, 1949 to present). FNO recorded temperatures at 0500/1700 (01UT/13UT) local standard time (LST), FCH at LST
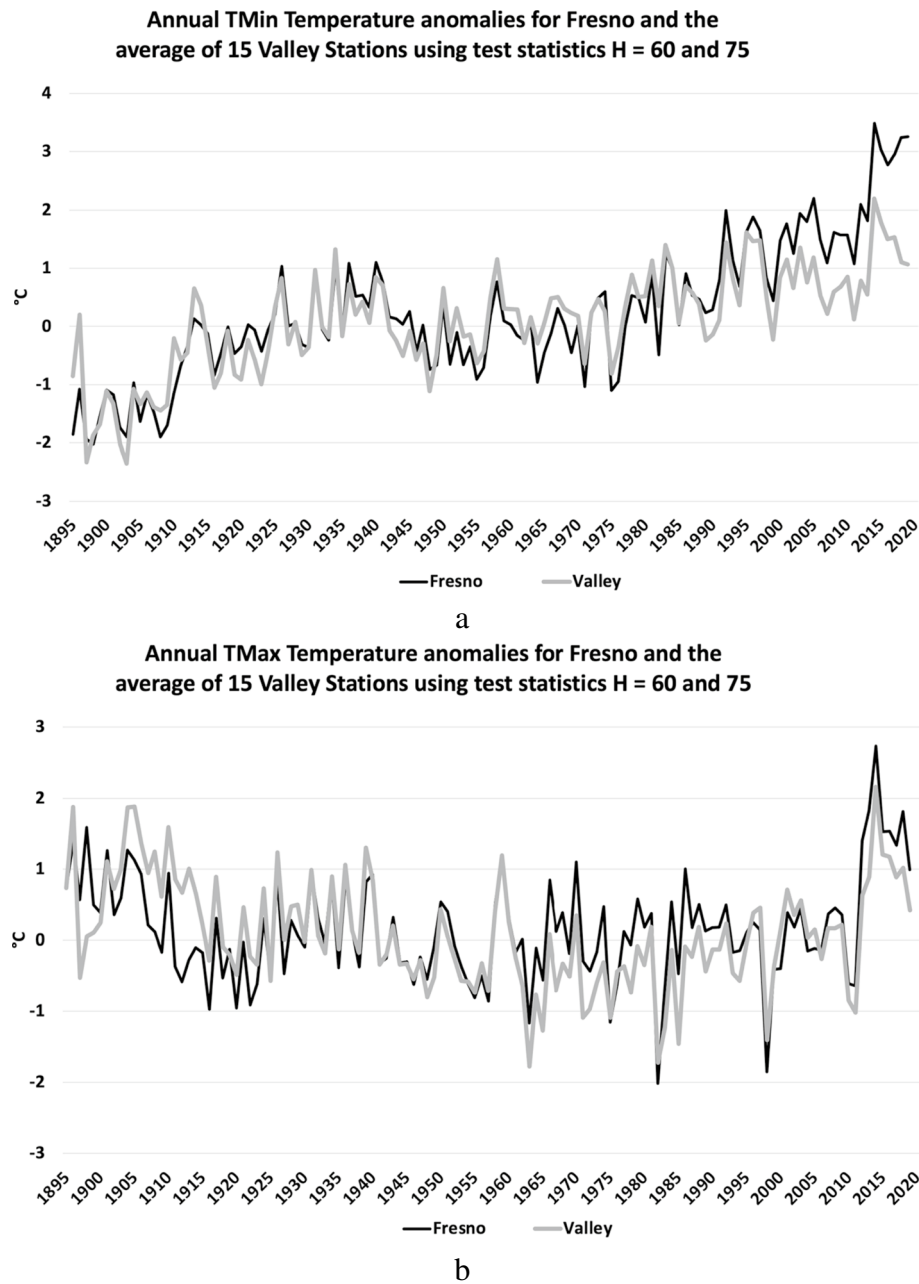
Fig. 6 **a** As in Fig. 4a with the adjusted time series as the mean of $H=60$ and $H=75$. **b** As in Fig. 5a but for TMax

0442/1642 (0042UT/1242UT), and FAT at 0400/1600 (00UT/12UT). The conversion to 00UT/12UT was accomplished through the interpolated difference calculated from

FAT temperatures, where all hours of the day in the last few decades were recorded. For example, in July, the average 12UT temperature was 0.63 °C warmer than 13UT at FAT;

thus, 0.63 °C was added to FNO and 0.44 °C to FCH readings to convert 0500 and 0442, respectively, to account for their later (slightly cooler) observation times. An additional shift was applied to account for the FNO's rooftop observations based on 6 years of overlapping data with FCH and for FCH's ground-level but urbanized site (Christy 2021).

The results are shown in Fig. 7 and resemble major aspects of the annual values in Figs. 6a, b. There is a period of relatively cooler readings in the middle of the 00UT (1600 LST) time series and a clear upward trend which accelerated after 1970 in 12UT (0400 LST). The adjusted July hourly time series produced 1895–2019 trend values of +0.42 (+0.02) °C decade^{-1} for 12UT (00UT). We note that this is an exploratory test because a single month cannot be expected to reproduce the values calculated for the annual time series in Fig. 6a, b—note, in particular, the annual values after 2012 are much higher than seen in July only and were due to anomalous heat in spring and fall months. In any case, this result provides a potentially new source of observations that may be useful for climate analysis due to its inherent advantages over the vagaries that affect observations of TMax and TMin.

6 Discussion

Through the years, surface temperature data have been recorded through times of substantial changes in instrumentation, surroundings, and practices. These changes often impact the record in ways that produce discontinuities for assessing long-term changes, not to mention the slowly evolving impacts of increasing infrastructure surrounding the stations. As noted earlier, many investigators have identified the sudden inhomogeneities in recent decades and attempted to produce adjustment algorithms so that a more useful time series unaffected by these changes could be studied.

That the local landscape has changed significantly is evident in Figs. 8 and 9. Figure 8 displays the US historical topographic maps, 7.5-min series, of the Fresno-Clovis area in 1922, the first such date maps were generated for this area. The map represents approximately 16 km N-S and 14 km E-W. The built-up infrastructure is depicted by the higher density of the streets and shows Fresno, in the lower left, as a rectangular array of NW–SE and NE-SW streets paralleling the railroad tracks. From there the city in 1922 had grown on the NE side of this rectangle and also due north with new streets now dominated by N-S and E-W directions. The population in 1920 was 45,000. The weather station up to 1949 was "downtown" (red dot) and the station since that time has been located at the Fresno Air Terminal (blue dot). The airport exposure is over dry, unvegetated ground, typical of the natural surface cover during most of the year.

Clovis, in 1920, was populated with 1150 people and occupied a small set of streets in the NW portion of the 1922 map. All told, less than 50,000 people lived in the area shown. The majority of the area was not urbanized, being utilized largely as farms with orchards, vineyards, and grain fields which the second author (Christy) well remembers from his early days in the 1950s.

Figure 9 is a 2021 areal photo of the same region now almost entirely urbanized (gray color), with no open areas except the groundwater recharge ponds NW of the airport runway (center-right) and the agricultural experimental farms of Fresno State University north of the ponds.

Fig. 7 Time series of monthly averaged hourly values observed at 00UT (1600 LST) and 12UT (0400 LST) for July
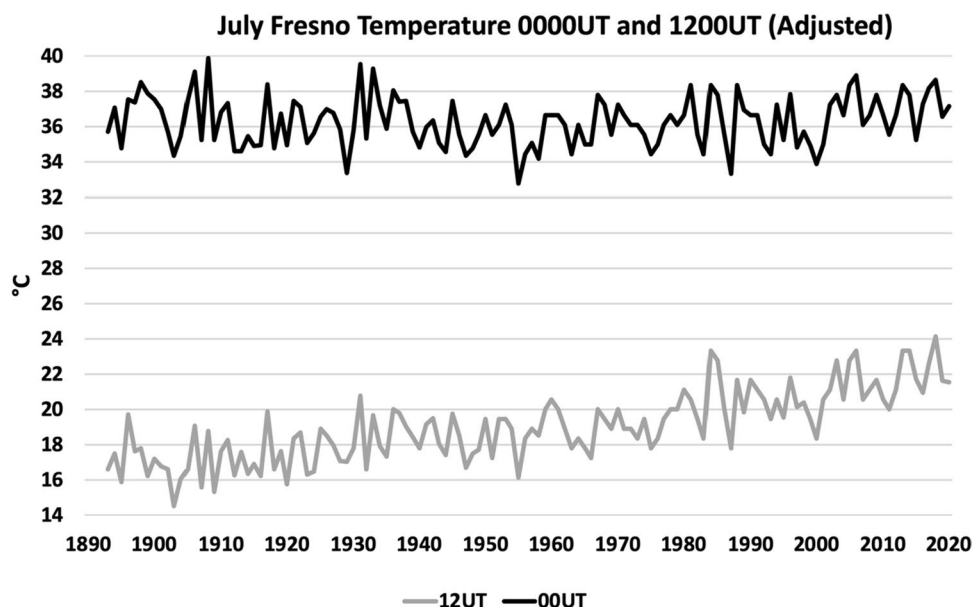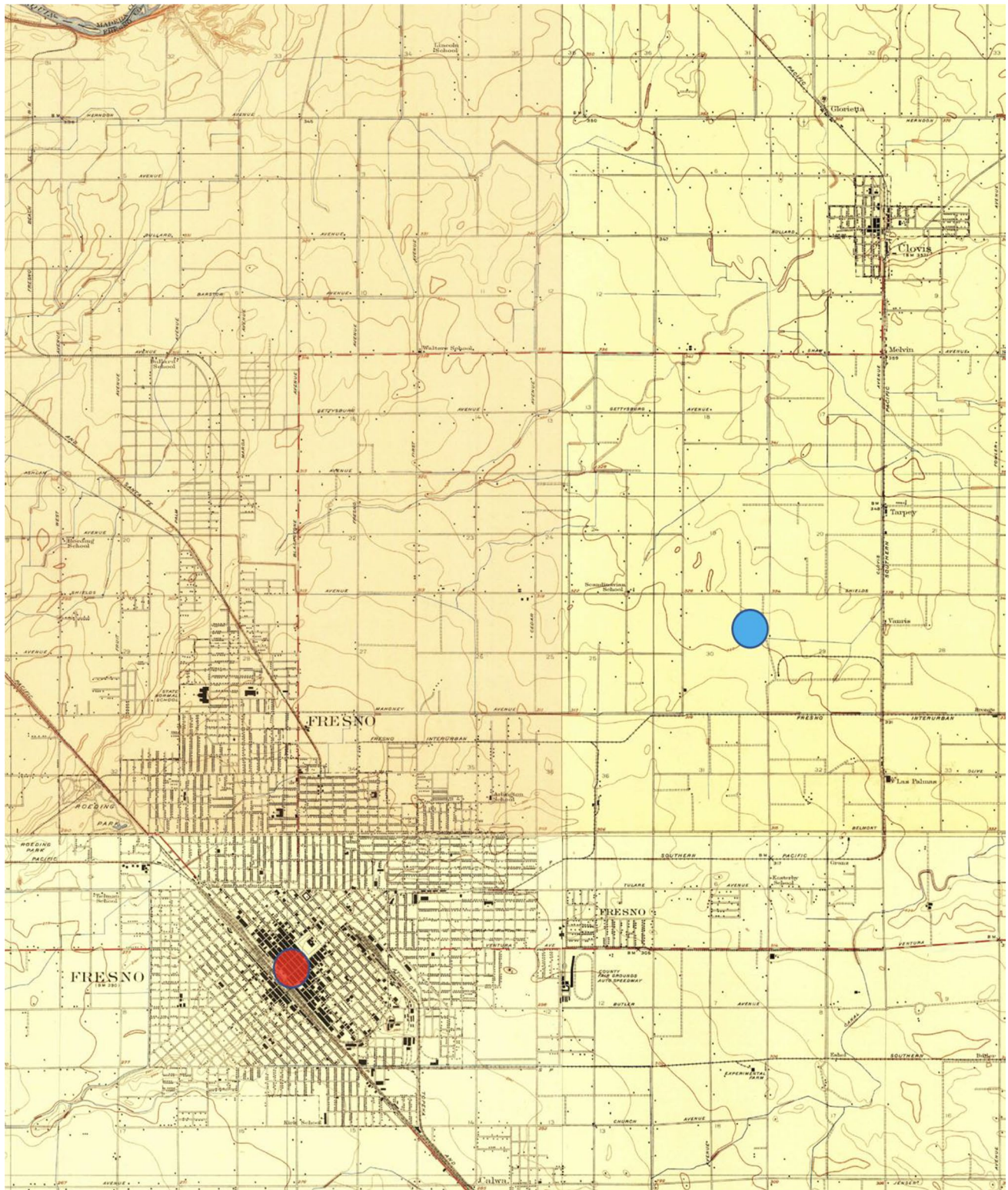
Fig. 8 Topographic map of the Fresno-Clovis area from 1922. Developed areas are indicated by closely spaced streets. The vertical extent is approximately 16 km, the horizontal extent is approximately 14 km. The center of the map is + 36.779 latitude, − 119.758 longitude
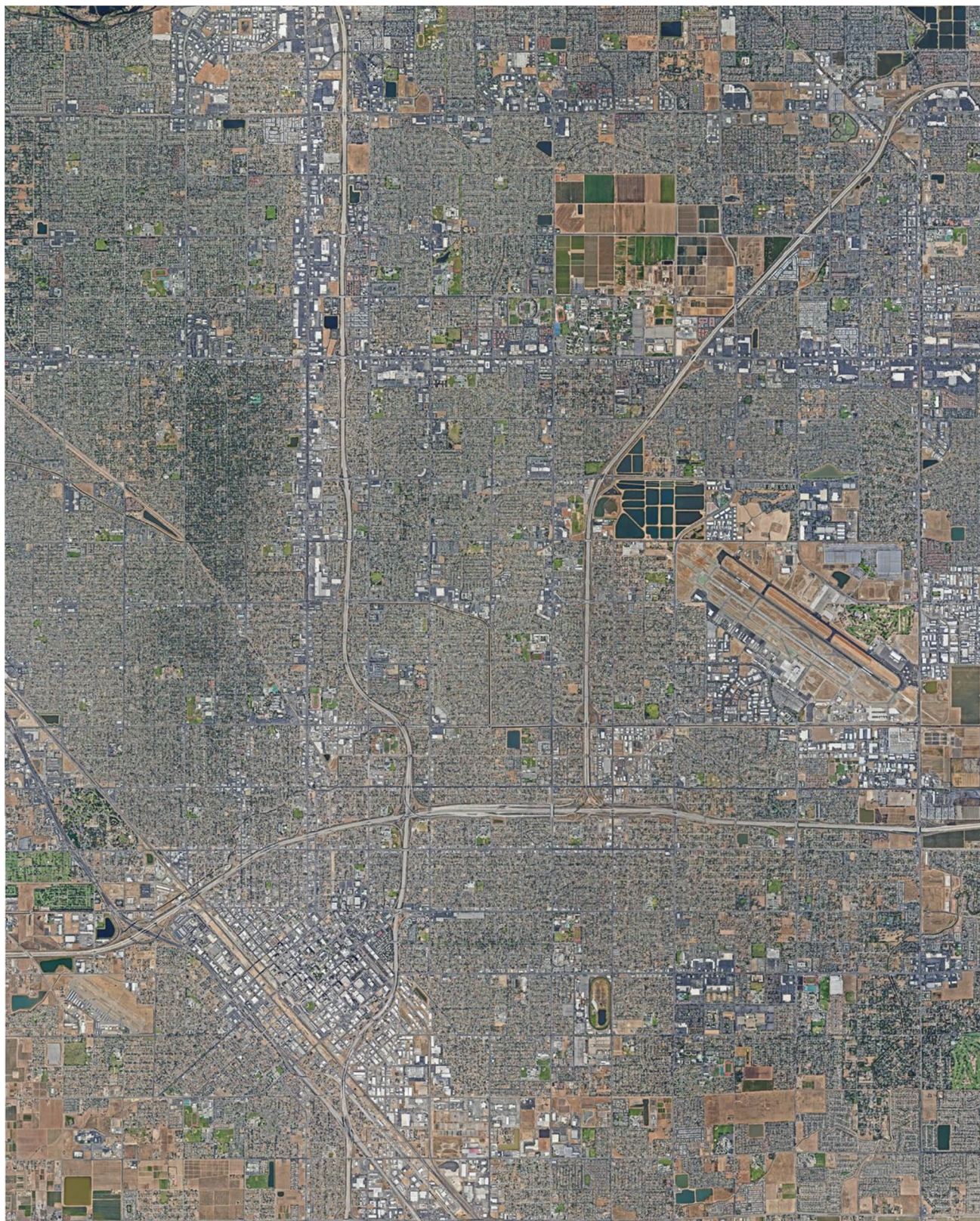
Fig. 9 A Google Earth image from 2021 of the Fresno-Clovis area matching that in Fig. 8. The current weather station is located between the runway and taxiway between the two cross-overs. The gray colors represent urbanization

It is obvious from these map representations that the surface influence on the local weather station at Fresno has substantially changed. Other evidence for change, in general, is found in photographs of individual stations which expose the instruments to unnatural perturbations in the resulting temperature measurement that could be characteristic of the other stations (e.g., Fig. 3 in Christy 2013a, b; Fig. 2 in Davey and Pielke, Sr. 2005; Figs. 1 and 2 Watts et al. 2015).

In this investigation we use the novel approach of defining the temporal extent of the metric as 5-day periods or pentads. In this way we reduce random noise of the daily metric for better statistical treatment of the analysis here performed. Requiring that all 5 days be present for the pentad calculation improves over typical studies which use monthly data but allow a number of missings to be acceptable. And, with 73 periods per year, the resolution of the annual cycle is more precisely determined for anomaly calculation.

We use a more-or-less classical approach for determining breakpoints, but have selected a procedure that focusses on a single station, Fresno, in a method that attempts to isolate its long-term change relative to the long-term changes of the remaining, much less urbanized stations nearby. The results indicate there is little difference between Fresno and non-Fresno stations for the metric of TMax. However, in the past 70 years of Fresno's rapid urban expansion (population of city limits increased from 91,700 in 1950 to 540,000 in 2022), the TMin change is three times that of the other stations. However, many of the comparator stations likely have experienced unnatural warming from their own growth. Evidence for this is the following population estimates (worldpopulationreview.com) for 1950 and 2022, respectively, for some of the stations analyzed; Madera (10.5 k, 67.1 k), Merced (15.3 k, 85.4 k), Porterville (6.9 k, 59.3 k), Reedley (4.1 k, 25.8 k), and Visalia (11.7 k, 138.1 k). With growth of five to over ten times in population, we may presume there is unnatural TMin warming in the non-Fresno stations, indicating Fresno's own unnatural warming is greater than just the difference between the two.

As explained in numerous studies, (e.g., Oke 1973; Oke et al. 1991; Nair et al. 2011; Fall et al. 2011; McNider et al. 2012; Christy et al. 2013; Scafetta 2021), the natural nocturnal cooling of urban areas (i.e., TMin) is inhibited by several factors which have much less influence on TMax. TMax is measured, in general, in mid-afternoon after the surface has been heated for several hours, creating deep vertical atmospheric mixing (1 to 2 km and more in depth) of the boundary layer which among other things allows the surface temperature to attain more of the character of a large mass of the atmosphere. Horizontal winds are of greater magnitude above the surface, so the vertical motions also mix these downward, expanding the sphere of influence of a surface thermometer even more. Thus, nearby stations with differing urbanization levels have greater affinity with one another in the afternoons through the deep vertical and horizontal mixing of the boundary layer air down to the surface.

The nocturnal boundary layer, in which TMin is usually observed, is altogether different. As the surface cools at night, the air becomes dense and separates (decouples) from the deep atmosphere above often being a layer only a few meters deep as a nighttime inversion occurs. This cold air does not represent the character of much warmer air above.

This shallow, cold layer is somewhat delicate and can be readily disturbed so that the much warmer air above mixes down to keep the surface temperature warmer than it would have been in the undisturbed state. Factors that create such disturbances include the presence of buildings as they disrupt the horizontal winds above the shallow, cold layer, forcing a mechanical mixing of the warmer air above down to the surface. Then, certain surfaces which absorb more heat than the natural ground cover will release that heat through the night, creating enough vertical mixing to prevent a full decoupling of the potentially cold surface from the warmer air above. Further, atmospheric constituents such as thermal-absorbing aerosols or greenhouse gases serve to retard the cooling rate of the surface, thus retarding the formation of the cold, decoupled surface layer (Nair et al. 2011). Note that in all of these cases, there is not an accumulation of more heat that affects the station, but a redistribution of heat (McNider et al. 2012). The end result of this is TMin experiences warming not found in pristine sites.

The results we see here for Fresno and the Valley are consistent with this well-established boundary-layer theory (McNider et al. 2012; Lin et al. 2016). For the broad atmospheric mass as observed in mid-afternoon (TMax), the values of the trends suggest there has been no significant change in the long-term temperature time series for Fresno ($+0.01$ °C decade^{-1}) nor for the Valley (-0.05 °C decade^{-1}) since 1895. NOAA/NCEI uses a different algorithm which, for example, incorporates stations further from Fresno for breakpoint detection and fewer nearby as is done here. Their resulting TMax (TMin) 125-year trends through 2019 are $+0.03$ ($+0.23$) °C decade^{-1}, being within error limits of our calculations and support the lack of atmospheric warming as indicated by TMax and significant warming of the urban nocturnal boundary layer at night (TMin).

However, since 1970, the NCEI analysis for Fresno diverges from trends in this study which calculated TMax (TMin) trends as $+0.27$ ($+0.63$) °C decade^{-1}, while those of NCEI are $+0.44$ ($+0.36$) °C decade^{-1} (accessed June 11, 2021, NCEI values tend to change as processing algorithms are updated). In the period after 1970, our method detected several breakpoints due to relocations, some to accommodate construction at the air terminal, as well as installation of new equipment. Note that our TMax analysis is corroborated by the combined Valley station time series for this period. Evidently, the combination of our breakpoint adjustments

versus those of NCEI in the last 50 years produces differing 50-year trends as indicated. There is clear evidence that nights have warmed significantly compared with days in this region in the past 50 years (e.g., Christy et al. 2006; Christy 2021 as well as Figs. 5a, b and 7 in our study), so this 50-year difference poses a question for further study. Indeed, we are at a loss to explain how the raw data can be adjusted to produce a TMax trend greater than TMin from 1970 to 2019. However, for the 125-year period, the differing breakpoint adjustments between ours and NCEI tended to essentially average out to near zero in their impact on the long-term trend.

The relative strong upward trend since 1970 for both Fresno and Valley in TMax (about $+0.3$ °C decade^{-1}) is likely a coincidence in that the beginning of this 50-year period was quite cool and that the worst regional drought in 130 years occurred near its end (2012–2016). The types of drought-causing stagnant high-pressure systems are not unusual and lead to both higher temperatures and lower precipitation. Indeed, in this climate, such drought periods of up to 100 years have occurred in the past when lakes in the neighboring Sierra Nevada Mountains receded so far and for so long that forests were established on the exposed lake bottoms. Today their drowned trunks are now submerged (in Lake Tahoe see Lindstrom, 1990, and in Fresno County see Morgan and Pomerleau, 2012). We note that the TMax trend for Fresno (Valley) for 1895–1970 was -0.08 (-0.20) °C decade^{-1}, so a rebound from this decline would be a statistically likely expectation.

Regarding the possible effect of increasing greenhouse gases, this rapid warming of TMin relative to TMax was not reproduced in the recent CMIP-6 climate model simulations. We accessed 28 CMIP-6 surface temperature time series for the conterminous US and their average result for 1970–2020 indicated the average TMax warmed insignificantly more than TMin ($+0.014$ °C decade^{-1}). Of the 28 models, most, (18) produced TMax trends greater than TMin which is opposite of this result found for Fresno. Thus, these simulations support the conclusion that Fresno's rapid rise of TMin relative to TMax is due to factors unrelated to large-scale forcing.

It is unfortunate that the two most common climate metrics observed have been the high and low temperatures in a 24-h period, TMax, TMin, and from them a computed average (TAvg). The ready availability of TAvg has led to it becoming the metric of choice even though it represents a convoluted indicator of temperature change over land. Studies such as ours should provide encouragement to the climate community to investigate TMax as a preferred long-term indicator of surface temperature change as it represents the deeper atmosphere and is less prone to the vagaries of the formation of the shallow nocturnal boundary layer and localized impacts of urbanization (McNider et al. 2012). As explained, TMin is extremely sensitive to the immediate landscape and its changing character over time and thus is contaminated too easily by these non-climatic factors.

Simply using Tmax, however, does not solve remaining problems associated with (a) urbanization as cities will still be hotter in the day than the countryside, (b) the time-of-observation bias, (c) location changes, and (d) instrumental upgrades. With that in mind, it is further recommended that long-term datasets of hourly temperatures be digitized and investigated for use as indicators of climate variability and change as they avoid some of the problems inherent with the extrema of daily temperatures.

7 Conclusion

In this investigation we demonstrate a method to improve the surface temperature values for climate studies now based on daily high and low temperatures recorded during 24-h periods. Using the temporal metric of 5-day averages (pentads), we generated daily high and low temperatures (TMax and TMin) for 16 stations in the San Joaquin Valley of California, USA, in an attempt to document the influence of infrastructure expansion around the largest city in the sample, Fresno, over the period 1895 to 2019. Applying a breakpoint detection and adjustment technique based on statistical significance and intercomparing all stations with each other, we created an adjusted time series of all stations in which these shifts relative to other stations were removed.

Comparing Fresno versus the average of the much less-urbanized stations, we detected a significant warming in Fresno TMin values, especially over the 50-year period of its largest growth 1970–2019. It is clear that TMin in Fresno has been impacted by urbanization with an estimated effect of at least $+0.4$ °C decade^{-1} as this is the amount our analysis suggests Fresno has warmed relative to the non-Fresno stations. Since there has likely been some urbanization impact on the non-Fresno stations, too, the value of $+0.4$ °C decade^{-1} should be considered a lower bound on the influence of surface development on TMin around Fresno.

This result is consistent with boundary-layer theory; TMin occurs in a shallow nocturnal boundary layer that may be systematically disturbed through time, resulting in an increasing tendency for warm air above to be mixed to the surface (i.e., impeding the usual decoupling of the surface cool layer from warm air above). Factors that cause such increased mixing include (a) buildings which disrupt the vertical wind profile to prevent the decoupling between the cool surface layer and the warmer air above, (b) increases in thermal absorbing constituents such as aerosols or greenhouse gases which retard radiational cooling of the surface, and (c) changes in surface type to those more conducive to increasing heat content (e.g., vegetation to concrete) and

keeping temperatures warmer through the night than otherwise would be the case. These all interact to redistribute the heat in the vertical profile from higher levels down to the surface, but do not increase the heat content in the total column by a meaningful amount. Since 1970, the trend of the time series of differences TMin minus TMax for Fresno was $+0.36$ °C decade^{-1}. The same metric for the non-Fresno stations was an insignificant -0.08 °C decade^{-1}, indicating again a warming rate for TMin unrelated to large-scale atmospheric forcing.

The overall TMax trend for Fresno and the non-Fresno stations during the period 1895–2019 was negligible, $+0.01$ and -0.05 °C decade^{-1}, respectively, with an estimated error of ±0.10 °C decade^{-1}. This is an indication that the larger scale climate variations have not experienced any unusual change since 1895 in this region. We also demonstrated that there may be value in hourly readings of temperature as they avoid some of the convoluted impacts on the metrics such as TMax and TMin.

Declarations

Ethics approval N/A. No human or animal subjects were investigated in this research and thus no consent forms are required.

Consent to participate Both authors consent to participate.

Consent for publication The University of Alabama in Huntsville consents to publish these findings.

Conflict of interest The authors declare no competing interests.

References

Christy JR (2001) When was the hottest summer? A state climatologist struggles for an answer. Bull Amer Meteor Soc 83:723–734

Christy JR, Norris WB, Redmond K, Gallo KP (2006) Methodology and results of calculating central California surface temperature trends: evidence of human-induced climate change? J Climate 19:548–563

Christy JR, Norris WB, McNider RT (2009) Surface temperature variations in East African and possible causes. J. Climate 22:3342–3356. https://doi.org/10.1175/2008JCLI2726.1

Christy JR (2013) Reply to "Comments on 'Searching for information in 133 years of California snowfall observations.'" J Hydrometeor 14:383–386. https://doi.org/10.1175/JHM-D-12-089.1

Christy JR (2013) Monthly temperature observations for Uganda. J Appl Meteor Clim 52:2363–2372. https://doi.org/10.1175/JAMC-D-13-012.1

Christy JR, McNider RT (2016) Time series construction of summer surface temperatures for Alabama, 1883–2014, and comparisons with tropospheric temperature and climate model simulations. J Appl Meteor Clim 55:811–826

Christy, J.R. (2021) *Is it getting hotter in Fresno … or not?* ISBN 9798714472664. 140pp. Available at Amazon.com.

Davey CA, Pielke RA Sr (2005) Microclimate exposures of surface-based weather stations - implications for the assessment of long-term temperature trends. Bull Amer Meteor Soc 86(4):497–504

Fall S, Watts A, Nielsen-Gammon J, Jones E, Niyogi D, Christy JR, Pielke RA Sr (2011) Analysis of the impacts of station exposure on the US Historical Climatology Network temperatures and temperature trends. J Geophys Res 116:D14120. https://doi.org/10.1029/2010JD015146

Griffith, B.D. and T.B. McKee (2000) Rooftop and ground standard temperatures: a comparison of physical differences. Climatology Report No. 00–2. Paper No. 694. Colorado State University.

Haimberger L (2007) Homogenization of radiosonde temperature time series using innovation statistics. J Climate 20:1377–1403

Karl TR, Williams CN Jr, Young PJ, Wendland WM (1986) A model to estimate the time of observation bias associated with monthly mean maximum, minimum and mean temperature for the United States. J Clim Appl Meteorol 25:145–160. https://doi.org/10.1175/1520-0450

Karl TR, Diaz HF, Kukla G (1988) Urbanization, its detection and effect in the United States climate record. J Clim 1:1099–1123. https://doi.org/10.1175/1520-042(1988)001

Karl TR, Jones PD, Knight RW, Kukla G, Plummer N, Razvayev V, Gallo KP, Lindseay J, Charlson JRJ, Peterson TC (1993) Asymmetric trends in surface temperature. Bull Am Meteorol Soc 74:1007–1023

Lin X, Pielke Sr. RA, Mahmood R, Fiebrich CA, Aiken R (2016) Observational evidence of temperature trends at two levels in the surface layer. Atmos. Chem. Phys., 16–827–841.

Lindstrom, S.G. (1990) Submerged tree stumps as indicators of Mid-Holocene aridity in the Lake Tahoe Basin, J. Calif. And Great Basin Anthro. 12(2).

McKitrick R, Michaels PJ (2007) Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data. J Geophys Res 112:D24S09. https://doi.org/10.1029/2007JD008465

McNider RT, Steeneveld GJ, Holtslag AAM, Pielke RA Sr, Mackaro S, Pour-Biazar A, Walters J, Nair U, Christy JR (2012) Response and sensitivity of the nocturnal boundary layer over land to added longwave radiative forcing. J Geophys Res 117:D14106. https://doi.org/10.1029/2012JD017578

Menne J.J, Williams CN Jr (2009) Homogenization of temperature series via pairwise comparisons. J Clim 22, https://doi.org/10.1175/2008JCLI2263.1

Morgan C, Pomerleau MM (2012) New evidence for extreme and persistent terminal medieval drought in California's Sierra Nevada. J Paleolimnol 47:7070–7713. https://doi.org/10.1007/s10933-012-9590-9

Nair UW, McNider R, Patadia R, Christopher SA, Fuller K (2011) Sensitivity of nocturnal boundary layer temperature to tropospheric aerosol surface radiative forcing under clear-sky conditions. J Geophys Res 116:Do2205. https://doi.org/10.1029/2010JD014068

Oke T (1973) City size and the urban heat island. Atmos Env 7:769–779

Oke TR, Johnson GT, Steyn DG, Watson ID (1991) Simulation of surface urban heat islands under 'ideal' conditions at night part 2: diagnosis of causation. Boundary-Layer Meteorol 56:339–358. https://doi.org/10.1007/BF00119211

Parker DE (1994) Effects of changing exposure of thermometers at land stations. Int J Climatol 14:1–31

Pielke RAJ Sr, Nielsen-Gammon C, Davey J, Angel O, Bliss N, Doesken MC, Fall S, Niyogi D, Gallo K, Hale R, Hubbard KG, Lin X, Li H, Raman S (2007) Documentation of uncertainties and biases associated with surface temperature measurement sites for climate change assessment. Bull Amer Meteor Soc 88(6):913–928

Scafetta N (2021) Detection of non-climatic biases in land surface temperature records by comparing climatic data and their model simulations. Clim Dyn. https://doi.org/10.1007/s00382-021-05626-x

Thorne PW, Willett KM, Allan RJ, Bojinski S, Christy JR, Fox N, Gilbert S, Jollieffe I, Kennedy JJ, Kent E, Tank AK, Lawrimore J, Parker DE, Rayner N, Simmons A, Song L, Stott PA, Trewin B (2011) Guiding the creation of a comprehensive surface temperature resource for twenty-first-century climate science. Bull Amer Meteor Soc 45:142. https://doi.org/10.1175/2011BAMS3124.1

Walters JT, McNider RT, Shi X, Norris WB, Christy JR (2007) Positive surface temperature feedback in the stable nocturnal boundary layer. Geophys Res Lett 34:L12709. https://doi.org/10.1029/2007GL029505

Watts, W., E. Jones, J. Nielsen-Gammon, J.R. Christy (2015) Comparison of temperature trends using an unperturbed subset of the U.S. historical climatology network. American Geophysical Union Fall Conference, San Francisco CA.

Williams CN, Menne MJ, Thorne PW (2012) Benchmarking the performance of pairwise homogenization of surface temperatures in the United States. J Geophys Res 117:D05116. https://doi.org/10.1029/2011JD016761