WHEN WAS THE HOTTEST SUMMER?

A State Climatologist Struggles for an Answer

BY JOHN R. CHRISTY

It takes painstaking detective work with original hand-written documents to answer seemingly straightforward questions about northern Alabama climate

Every state climatologist receives numerous and diverse requests for climate information. We often work closely with local and state development offices to provide data for potential businesses that must factor in the costs and benefits of climate impacts on their operations. For example, a request came from an anonymous corporation wanting to establish a billion dollar business in a certain Alabama county, but they needed 24 hours per day/7 days per week (24/7) access to highway transportation. We provided them with an estimate of highway closures due to ice and snow, which was viewed favorably by the corporation. Another request, from a European chocolate maker, required information on the magnitude and duration of high dewpoint temperatures. The answer proved to be a significant challenge for candy making in Alabama.

Most requests, however, come from the media who ask for the most accurate and up-to-date perspective on some type of recent or developing climate anomaly; that is, has it been this hot before?

Hampering our ability to answer these requests is the lack of consistently observed weather variables over the timescale of a century so that probabilities of events that occur on the order of a few per century may be estimated with some confidence. Customers expect that since Alabama's Office of the State Climatologist is located in Huntsville, Alabama, their requests that deal with temperature in the local area of north Alabama should be easily accommodated. Unfortunately, this is not the case.

Several stations have operated within a relatively small region around Huntsville in extreme north Alabama and their data are nominally available to try to answer the questions posed. Most of these stations were cooperative sites manned by volunteers (Sidebar 1). None of these stations, however, has operated continuously since the first was established in the late nineteenth century, and those with the longest records are now closed.[1] So, in answering the question, When was the hottest summer in Huntsville, Alabama? one is reluctant to offer any answers since rigorous station-by-station intercomparisons for the period of record have not been performed.

AFFILIATION: CHRISTY—University of Alabama in Huntsville, Huntsville, Alabama
CORRESPONDING AUTHOR: Dr. John R. Christy, Alabama State Climatologist, Earth System Science Center, University of Alabama in Huntsville, Huntsville, AL 35899
E-mail: christy@nsstc.uah.edu

In final form 4 December 2001.

---

[1] Monthly averages of Huntsville temperature data from 1829 to 1842 have been located, but what the values actually represent is unknown. Data from forts exist for several Alabama sites prior to 1880, though none were in northern Alabama.

Typically, when responding to such questions, we in the office access an online data archive of observations taken at Huntsville's National Weather Service (NWS) office presently located at the Huntsville/Decatur International Airport. Unfortunately, online data are only available from 1958, the year the current station was commissioned. Yet, even these data are problematic. In 1958, the original site was essentially within the city's urban area then was moved 18 km west in 1967 to the newly opened airport which was surrounded by cropland and forest. Since 1967, considerable development has occurred around the airport's east and north sides. A second major event occurred with the installation of an automated observing system in 1994, which required the instruments to be moved from a site near the terminal building, surrounded by asphalt, out to the edge of the west runway with no urbanization nearby. Thus data from the local NWS office are inhomogeneous and of relatively short duration. This type of problem exists throughout the country.

CLIMATOLOGISTS' UNSUNG HEROES

The people who braved the elements to take daily weather observations since the nineteenth century have proven to be true foot soldiers of science. These people operate what are called "coop" stations because these sites are operated by "cooperative" observers. As the twenty-first century begins, we are able to utilize the results of their unpaid, dedicated service to answer important scientific questions such as how has the frequency of flooding rains or occurrence of high temperatures changed? Because most of these stations were sited to observe agricultural development (i.e., first frost, cotton bursts) they have largely remained in rural, relatively unchanged local environments.

These volunteers persevered through many problems, but some things are beyond their control. The observer in Alco, Alabama, apologized for lack of communication in Oct 1897 because he was prevented by the Health Department, which had him quarantined during a yellow fever outbreak. Thomas W. Carter, an observer for 33 years in Madison, Alabama, recorded that he was struck with the flu in the great 1918 influenza epidemic and unable to perform his duties for 16 days in Dec of that year. On 24 Jul 1924 Mr. Carter also dealt with a microburst or tornado that "demolished thermometer shelter and broke thermometers." No measurements were taken for 9 days.

What makes these handwritten records especially valuable is the detailed comments that were used to document small changes in the way the observations were taken. This information is often called metadata and includes, for example, statements that the instrument shelter was moved 20 feet for better exposure or that a thermometer was malfunctioning and a new one installed. Notes like these, which are not available in digitized form, must be individually read; however, they are indispensable in helping determine inhomogeneity points in the data for potential corrective action.

DATA AND METHOD. To answer the question in a scientifically defensible manner I calculated for every station near Huntsville the summer monthly average [June, July, August (JJA)] of the daily maximum temperatures. Then, I applied a deliberately simple technique to produce a complete, homogenized time series to attempt in answering the question.

To construct the time series, I initially used 13 stations within a 45-km radius of the current NWS site. These are referred to as the "Near" stations. In any dataset reconstruction technique, however, it is important to provide some type of independent check of the results. I then selected 5 stations outside this radius (50–85 km or "Far" stations) and again applied the merging technique, creating an independent set of data (Fig. 1). The stations in the Near dataset were all within 40-m elevation of the present Huntsville NWS office. The data begin in 1893 when the first reference of the use of the Cotton Region Shelter was documented for these stations.

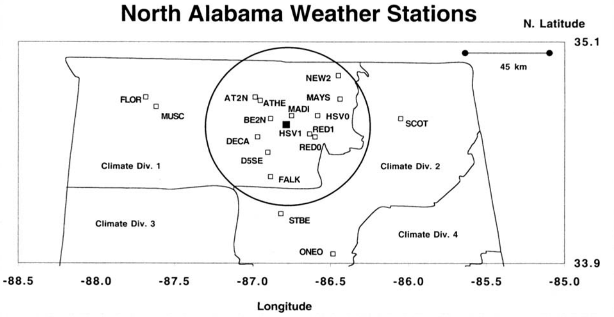

FIG. 1. Map of the stations used in this study. Locations are only approximate as some have moved as much as 18 km during their operation. The circle represents the 45-km radius around the current NWS office. The 13 Near stations are those within the circle; the 5 Far stations lie outside the circle.

Observations and metadata were obtained for these 18 stations from (i) National Climatic Data Center's (NCDC's) *Climatological Data* publications on paper and microfiche, (ii) the Southeastern Regional Climate Center Web site, (iii) original handwritten paper forms in Alabama's Office of the State Climatologist, and (iv) U.S. Army observations at Redstone Arsenal.

The idea behind the homogenization technique is to identify points in time in each station's record at which a change of some sort occurred. These are called segment break points. A single station may have a number of segment break points so that its entire record becomes a set of segments each of which requires some adjustment to make a homogeneous time series. Initially, segment break points were identified in every case when one of the following situations occurred: (i) a station move, (ii) a change in time of observation, and (iii) a clear indication of instrument change.

Much of this break point information was gathered through reading the original handwritten forms, page by page, on which observers or government validators made notes in the margins. This significant human intervention was required since much of the metadata information (i.e., handwritten notes) on each form is nowhere digitized for easy access.[2] This tedious and imperfect process suggests that it is certainly probable that all break point events were not recorded.

This procedure converted the time series of the 13 Near stations into a total of 46 individual segments (Table 1). However, 2 other break points of unknown

TABLE 1. List of homogeneous segments from 14 stations contained within the Near network. The full coop (see sidebar 1) ID number should include the state prefix 01 (e.g., 01-0390 for Athens). The reasons for designating each segment break relative to the previous segment (immediately above) are "I" instrument change, "L" location change, "T" time of observation change, and "U" unknown. The magnitude of the bias vector b (see text) indicates the value to be subtracted from the raw temperatures of the segment so as to be consistent with segment 32, which is the current NWS office in Huntsville.

Segment (s)	Station	Coop ID	First entry	Last entry	Reason	No. months observed	No. months comparison	Bias relative to HSV1 |b|
1	ATHE	0390	Jun 1991	Aug 1994		10	31	−1.78
2	ATHE	0390	Jun 1995	Aug 1997	U	9	27	−0.28
3	ATHE	0390	Jun 1998	Aug 2000	T	9	27	−0.40
4	ATN2	0395	Jun 1956	Aug 1956		3	14	0.29
5	ATN2	0395	Jun 1957	Aug 1962	T	18	126	0.09
6	ATN2	0395	Jun 1963	Aug 1986	L	67	371	0.53
7	ATN2	0395	Jun 1987	Aug 1988	I, T	5	20	0.47
8	ATN2	0395	Jun 1989	Aug 1990	L	6	23	0.37
9	BE2N	0655	Jun 1950	Aug 1962		39	200	0.74
10	BE2N	0655	Jun 1963	Aug 1974	L, T	36	259	0.49
11	BE2N	0655	Jun 1975	Jun 1994	T	57	203	0.34
12	BE2N	0655	Jul 1994	Aug 2000	L	20	59	0.66
13	DECA	2207	Jun 1893	Aug 1903		33	64	0.99
14	DECA	2207	Jun 1904	Aug 1919	T	48	89	0.31
15	DECA	2207	Jun 1920	Aug 1927	T	24	44	−1.06
16	DECA	2207	Jun 1928	Aug 1933	T	18	34	−0.47
17	DECA	2207	Jun 1934	Aug 1954	T	63	145	−0.21
18	DECA	2207	Jun 1955	Aug 1956	L, T	6	26	0.02
19	DECA	2207	Jun 1957	Aug 1957	L	3	18	0.07

---

[2] During weekends and evenings, the author performed the task of reading all of the original data forms to document changes for this study.

Table 1. Continued.

Segment (s)	Station	Coop ID	First entry	Last entry	Reason	No. months observed	No. months comparison	Bias relative to HSV1 |b|
20	DECA	2207	Jun 1958	Aug 1959	L	6	41	0.22
21	DECA	2207	Jun 1960	Aug 1968	L	27	210	0.59
22	D5SE	2209	Jun 1999	Aug 1999		3	12	–0.97
23	FALK	2840	Aug 1956	Jul 1963		21	146	0.42
24	FALK	2840	Jun 1964	Aug 1970	L	21	157	0.32
25	FALK	2840	Jun 1971	Aug 1971	L	3	20	0.88
26	FALK	2840	Jun 1972	Jun 1992	L	60	245	0.44
27	HSV0	4064	Jun 1940	Aug 1940		3	9	0.39
28	HSV0	4064	Jun 1941	Aug 1945	L	15	30	0.02
29	HSV0	4064	Jun 1946	Jun 1954	L	25	63	–0.21
30	HSV1	4068	Jun 1959	Aug 1967		27	209	0.66
31	HSV1	4068	Jun 1968	Jul 1994	L	80	349	0.30
32	HSV1	4068	Aug 1994	Aug 2000	L, I	19	56	0.00
33	MADI	4976	Jun 1894	Aug 1906		39	76	0.45
34	MADI	4976	Jun 1907	Aug 1911	L	15	30	–0.08
35	MADI	4976	Jun 1913	Aug 1913	L	3	5	1.34
36	MADI	4976	Jun 1914	Aug 1915	T	6	12	0.73
37	MADI	4976	Jun 1917	Aug 1948	L	92	185	0.09
38	MADI	4976	Jun 1949	Aug 1949	T	3	6	0.44
39	MADI	4976	Jun 1950	Aug 1950	L, T	3	9	0.11
40	MADI	4976	Jun 1951	Jun 1962	L, T	34	175	0.33
41	MADI	4976	Jun 1963	Aug 1974	L, T	32	236	0.61
42	MAYS	—	Jun 1893	Aug 1893		6	9	–0.49
43	NEW2	5867	Jun 1959	Aug 1961		8	61	0.05
44	NEW2	5867	Jun 1962	Aug 1967	L	18	140	–0.41
45	NEW2	5867	Jun 1968	Jun 1975	L	19	127	0.34
46	RED0	6833	Aug 1954	Aug 1969		45	314	–0.15
47	RED0	6833	Jun 1970	Aug 1974	U	15	96	1.05
48	RED1	—	Jun 1981	Aug 1999		54	190	0.83
49	FLOR	2971	Jun 1893	Aug 1912		64	116	0.03
50	FLOR	2971	Jun 1913	Jul 1914	L	4	8	0.52
51	FLOR	2971	Aug 1914	Jun 1921	L, T	20	34	0.38
52	FLOR	2971	Jul 1921	Aug 1934	L, T	39	77	–0.47
53	FLOR	2971	Jun 1935	Aug 1935	T	3	6	–0.71
54	FLOR	2971	Jun 1936	Aug 1941	L	16	34	0.03
55	FLOR	2971	Jul 1954	Jul 1956	L	7	29	1.23

Table 1. Continued.

Segment (s)	Station	Coop ID	First entry	Last entry	Reason	No. months observed	No. months comparison	Bias relative to HSV1 |b|
56	FLOR	2971	Aug 1956	Aug 1960	L	13	87	0.43
57	FLOR	2971	Jun 1962	Aug 1963	L	6	44	1.67
58	FLOR	2971	Jun 1964	Aug 1965	T	6	48	1.87
59	FLOR	2971	Jun 1966	Jun 1979	T	36	225	1.36

origin were evident after the first reconstruction, so these were removed also (Sidebar 2).

I pause here to note that other methods to account for inhomogeneities have been devised and applied to surface data such as these. Perhaps the standard for homogenizing U.S. temperature data has been developed by scientists at National Climatic Data Center (Peterson et al. 1998b). These techniques determine adjustments for time of observation (Karl et al. 1986; Peterson et al. 1999), station history (i.e., location; Easterling et al. 1996), transition from liquid-in-glass to electronic thermometers (Quayle et al. 1991), missing data (Karl et al. 1990), and urbanization (Karl et al. 1988). Independent, statistical tests have also been utilized to detect inhomogeneities without station histories (Easterling and Peterson 1995; Peterson et al. 1998a). More recently, Hansen et al. (2001) have combined some of these adjustments with their own assessments that account for missing data and urbanization.

Because these other methods are intended to produce homogeneous time series for large geographic averages, the adjustments are generally systematized for the entire domain. Our study focuses on a very small subset of stations and thus I elected to determine individual, station-specific segment adjustments. My reasoning is that each station has unique factors that may depart from a systematic relationship determined from a large sample. For example, a station may change the time of observation from morning to evening in a variety of ways. Some stations indicate a change in the time as sunrise to sunset. Another may record a 9:00 A.M. to 5:30 P.M. change while yet another moves from 6:00 A.M. to 9:00 P.M. All are morning-to-evening shifts, but likely to be unsystematic.

MYSTERIOUS CLIMATE CHANGE?

After a preliminary intercomparison of the 46 segments, 2 additional, significant, and unexplained breaks or shifts in temperature were discovered. These are the kind of situations that generate confusion in constructions of this type. One example was found in Athens (ATHE). When compared with the average of 3 other co-observing stations, ATHE indicated a highly significant, permanent 1.5°C shift to warmer temperatures between 1994 and 1995 (Fig. SB2). No information from official records or from interviews with observers at the water treatment plant revealed a change in the site or instruments between Aug 1994 and Jun 1995. The standard error of the mean for the average of the other 3 stations on either side of the break was a very small 0.12°C, thus verifying that the 1.5°C shift was spurious. A similar analysis indicated a break in the Redstone Arsenal (RED0) site between 1969 and 1970. These 2 additional breaks were inserted, bringing the total number of segments to 48 for the 13 Near stations. Other, less dramatic, corrections were applied to MADI 1920, SCOT 1944, and ONEO 2000.

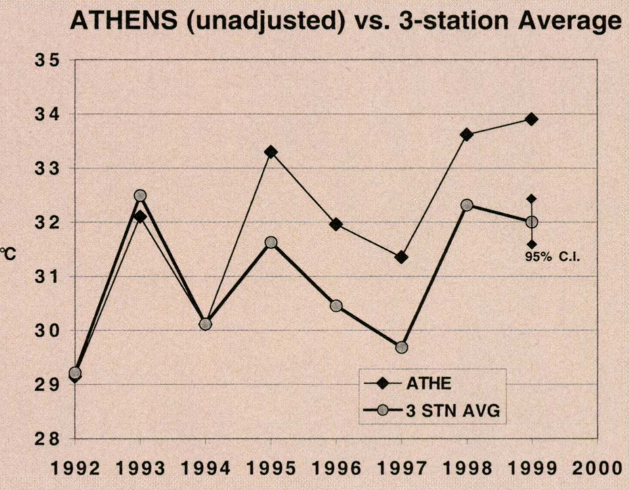

Fig. SB2. Time series of JJA daily maximum temperatures for the average of 3 nearby stations (shaded circles) and ATHE (solid diamonds) for 1992–99. The 95% confidence interval (C.I.) for the 3 stations is shown.

METHOD USED TO CALCULATE TABLE 1

The segment biases were calculated by a cumulative/iterative technique in which I constructed a bias vector (b) which related the bias of each segment to every other segment. Because many segments do not have direct overlaps with other segments, the bias for these must be calculated indirectly. This is similar to a ranking system in sporting events that must accommodate the fact many teams have not played in head-to-head competition. In this method, any segment may be chosen as a reference segment (i.e., bias of zero). I arbitrarily chose the latest segment of HSV1, the NWS office, as the reference.

A matrix was calculated in which the difference ($\delta_{i,j}$) of each segment (s_i) relative to every other segment (s_j, where overlapping data are available) was represented. A schematic example is shown in Table SB3. Note that $\delta_{i,j} = -\delta_{j,i}$ and that many entries are blank, indicating there is no overlapping period of observation between the two segments.

I selected as the initial bias vector (b) a first guess, which was the segment (or row) that contained the largest number of coincident observations with other segments, and thus the largest number of directly computed biases (i.e., head-to-head competition). In this case, segment 9 (BE2N) provided direct bias calculations with 27 other segments, more than any other, and was selected as the first guess. Segment 21 (DECA) was the second vector selected.

Suppose row s_6 in the sample matrix of Table SB3 contained the highest count of δs across the row, s_6 would be selected as the first guess for b (which in our actual case would be segment 9 of BE2N.) Suppose s_1 had the second highest count of δs (segment 21 in the actual case). A single difference-value computed between s_6 and s_1 was determined as the weighted-average difference of the δs common to both (e.g., from s_2 and s_4 above calculate the weighted average of $\delta_{6,2}$ minus $\delta_{1,2}$ and $\delta_{6,4}$ minus $\delta_{1,4}$). This is analogous to a system in sports by which two teams who have not played head-to-head are ranked through games played against several common opponents. In this example, had there been a common period of overlap between s_1 and s_6, the entry $\delta_{1,6}$ would also be employed to determine the mean difference between the s_1 and s_6. The mean difference between the two vectors is a weighted value based on the number of months observed in common.[SB1] Once the mean difference between s_1 and s_6 was determined and removed from all δ's in s_1, the two vectors were combined to form the second iteration of b. The vector b now became the base against which the next s vector with the third most counts of ä was compared. This procedure was followed until the final segment was merged into b and a value for each segment was computed.

Each segment bias (i.e., the elements of b) was determined as a mean of differences computed from many intercomparisons. The standard error of the mean for each entry in b varies due to the wide variation in the number of segment months available for each overlap for each segment but in general is less than 0.1°C.

TABLE SB3. Schematic representation of the bias matrix in which each entry $\delta_{i,j}$ is the mean difference between homogeneous segments *i* and *j*. Each row and column represents the relative bias between the homogeneous segments denoted as s.

	s_1	s_2	s_3	s_4	s_5	s_6	s_7	......	s_N
s_1	0	$\delta_{1,2}$	—	$\delta_{1,4}$	$\delta_{1,5}$	—	—		—
s_2	$\delta_{2,1}$	0				$\delta_{2,6}$			
s_3	—		0			—			
s_4	$\delta_{4,1}$			0		$\delta_{4,6}$			
s_5	$\delta_{5,1}$				0	—			
s_6	—	$\delta_{6,2}$	—	$\delta_{6,4}$	—	0	$\delta_{6,7}$		—
s_7	—					$\delta_{7,6}$	0		
·									
·									
·									
s_N	—					—			0

[SB1] Other weighting schemes were tested including the square root of the number of months, the inverse of the standard deviation of the differences, and no preferential weighting at all. Very little impact was observed.

Additionally, such observational shifts are often accompanied by another change in the observational context, such as location or observer, which only confuses the issue. I treated every change as a unique break point in the individual station's time series and determined the magnitude of the bias by empirical comparisons with overlapping data from other stations.

To enhance confidence in the pre-1940 data, which has only 2 stations in the Near network, I selected the "best" station from the Far network—Florence, Alabama (FLOR)—which provides data continuously from 1893 to 1940 (the period in need of more data) and data from 1954 to 1979, in a total of 11 segments. So, for the Near network average JJA maximum temperature, we utilize 59 segments determined from the 14 stations. My assumptions are as follows in applying the merging technique for the 59 segments:

- Each segment is a homogeneous time series and differs from other segments by a simple bias.
- Spurious trends over a segment are small and random.
- The calculated segment bias is the same for each of the months June, July, and August.
- Remaining time-varying differences between stations are due to random fluctuations in the natural variability of temperatures over the region and to random errors in the measurement systems (including the effect of missing observations).

To merge the individual segments I calculated and removed biases determined during periods of observation overlapping with other segments. This exercise, however, is an overspecified problem because at a particular segment breakpoint generally there are overlapping observations from many stations so multiple bias solutions are possible. All of these segments in turn have differing overlapping periods with other segments so that it is not possible to solve for a unique set of segment biases, in which all differences are perfectly consistent.

I chose to determine the segment biases by a cumulative/iterative technique in which I constructed a bias vector that related the bias of every segment to a single, reference segment (and thus to every other segment). This is similar to a ranking system in sporting events that must accommodate the fact many teams have not played in head-to-head competition (Sidebar 3). I arbitrarily chose the latest segment of HSV1, the NWS office, as the reference.

The standard errors for individual summer averages ranges from 0.01° to 0.86°C (1912 when only 2 stations reported). The mean standard error was 0.21°C (for a 95% confidence interval of approximately 0.4°C). The standard errors for individual segment biases were smaller because there were many more values (in monthly units) included in the calculation for each segment bias than would be found in the computation of the temperature of a single year. Generally speaking, individual segment biases are known to within ±0.1°C by this technique.

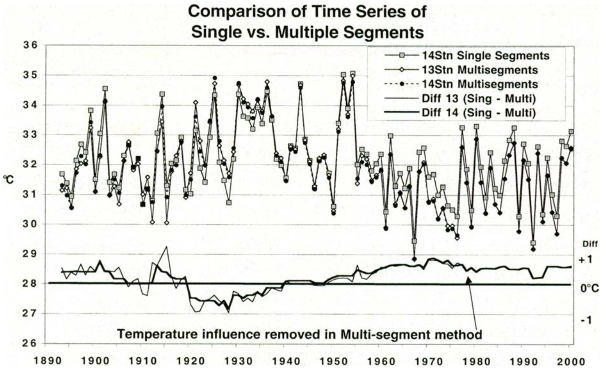

FIG. 2. Time series of adjusted JJA values of the multisegmented station records for 13 Near stations, multisegmented 14 stations (Near plus FLOR), single-segmented station records for 14 stations. (lower) The difference between the 13-station adjusted multisegmented time series and the single segmented time series (thin solid line) and the difference between the 14-station adjusted multisegmented time series and the single-segmented time series (thick solid line).

In Fig. 2 we show the comparison between the adjusted time series for the 13 Near stations and that of the Near stations plus FLOR (i.e., 14 stations) containing 48 and 59 segments, respectively (both noted as "multiple" segments). A third time series is included which assumes no discontinuities in each station record, or equivalently, that each station's complete record is a single homogeneous segment (noted as "single" segments). The interstation single-segment biases were calculated as before and the time series merged. The difference between this unadjusted time series and the 13- and 14-station networks is shown beneath. It is apparent that with the addition of FLOR, the adjustments are a bit smoother in the early part of the record for the multisegment method. Also, the addition of FLOR has virtually no impact on the overall 108-yr trend. Indeed in one additional computation, the values of Oneonta, Alabama (ONEO), were included so that 4 stations operating prior to 1940 were available, giving a total of 15 stations in the network. The 108-yr trend differences between the 13-, 14-, and 15-station networks were less than 0.005°C decade^{-1}, perhaps signifying some level of robustness in the technique.

SMALL LOCAL CHANGES MAY CAUSE LARGE TEMPERATURE EFFECTS

In looking at the individual bias corrections I draw attention to two stations, Huntsville (HSV1) and Athens 2 (ATN2, Table 1). The results tell us that moving the Huntsville NWS office (HSV1) from the city to a rural area (segments 30 to 31, Table 1) introduced a cooling of 0.36°C (0.66° minus 0.30°). Then, the installation of the ASOS instrumentation in 1994 contributed another 0.30°C cooling (segments 31 to 32). Thus, taken on its own without adjustment, the NWS data would show a spurious cooling trend over the period 1958 to 2000 as revealed by these results. Note too that the present site (segment 32) is indistinguishable from the original in-town airport site to a level of 0.02°C during 1940–45 (segment 28) when Huntsville was much smaller.

Another interesting result is evident for ATN2. The observer had noted that he did not believe the exposure of the instrument shelter was proper. On 5 Apr 1963 the station was moved 7 m to the north, a change noted in the original handwritten observations form and on WB-531-1, a form submitted for any change at the station. As shown in Table 1, this 7-m move at ATH2 (segment 5 to 6) is associated with a discontinuity of +0.44°C (0.53° minus 0.09°). The segment bias difference of +0.44°C (suggesting a spurious warming) is highly significant as segments 5 and 6 utilized 126 and 371 segment months, respectively, for their bias determination. Given the observation that most stations throughout their histories have experienced more rather than less development in their immediate vicinity, it is likely that even rural stations such as ATN2 contain spurious trends. These two examples show how spurious cooling or warming shifts may appear in the data.

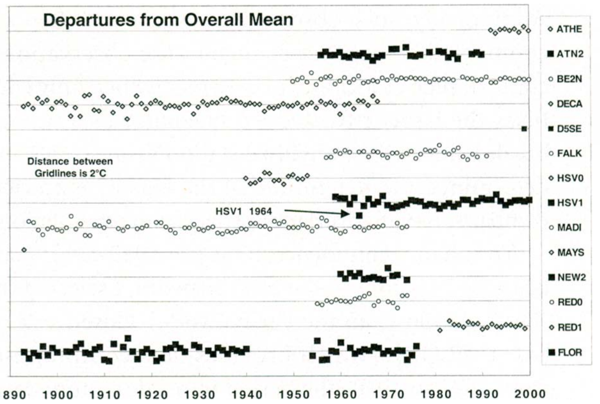

FIG. 3. Time series of JJA temperature departures from the grand mean for each station after multisegment adjustments have been applied. Note that distance between horizontal grid lines is 2°C.

The general result suggests that between 1893 and the 1920s stations experienced spurious cooling (Fig. 2, lower). From the 1920s to about 1970, spurious warming apparently occurred. Since 1970, temperatures have required little adjustment as determined by this method. It is important to note that virtually all of these stations would be considered rural in any reconstruction. The trend of the adjustments (which is removed) is +0.06°C decade^{-1}, suggesting that unadjusted JJA maximum temperatures from mainly rural stations may have a slight spurious warming due to the combination of factors indicated earlier.

In Fig. 3 we show the departures from the overall JJA mean for each adjusted station. Standard deviations range from 0.16°C [Athens, Alabama (ATHE)] to 0.39°C (FLOR, to be expected as this station is farthest from the centroid of the region). Station BE2N is exceptionally stable especially after 1970 with a standard deviation of 0.12°C in the last 30+ yr. This station has resided within the same small pasture since commissioned in 1950 by the Auburn University Agricultural Extension Service. Virtually no changes have occurred in the immediate vicinity. Two other stations are of interest regarding the impact of local changes on the temperature history (Sidebar 4). Other odd results, such as the Huntsville National Weather Service office relatively cold 1964 summer (Fig. 3, arrow), are discussed in Sidebar 5. A comparison between the Near and Far networks is discussed in Sidebar 6.

There are few avenues available to assess the confidence of the trend estimate. I am encouraged however that the differences in 108-yr trends of the 13-, 14-, and 15-station Near networks, having 2, 3, and 4 stations, respectively, in the 1893–1940 period, varied by less than 0.005 C decade^{-1}. However, there may be a set of peculiar coincidences in this matrix method that conspire to create errors in individual values and thus the overall trend.

DISCUSSION. The completed time series of the JJA mean of the reconstructed local area average of daily maximum temperatures is displayed in Fig. 4 along with the individually adjusted station values. With this result, I would respond to the question posed in the article's title that the warmest year of summer afternoons in Huntsville is probably 1925 (34.91°C), the year of the famous Scopes "Monkey" Trial in nearby Dayton, Tennessee. However, 1952 (34.75°C), 1930 (34.74°C), 1954 (34.72°C), and 1936 and 1943 (both at 34.59°C) are within the ±0.4°C margin of error. The

ODD RESULTS IN A FEW STATIONS

A few individual values for stations were highly suspicious and were examined closely to understand the nature of the reports.

MADI 1920. A note on the Mar 1921 MADI monthly report form states, "Max thermometer was a retreater." In these thermometers, the index in the thermometer tube should ride to the maximum value and remain there until reset. However, as the liquid contracted when the temperature fell, a "retreater" index would slip back from the maximum level, rendering spuriously cold readings. (Retreaters were not uncommon in the early years.) A replacement was provided in Mar 1921. A comparison with DECA and FLOR clearly shows the problem apparently had been ongoing since 1920. Subtracting the 1919 JJA value from 1918 gives for MADI, DECA, and FLOR –1.59°, –2.17°, and –1.50°C, respectively, indicating good agreement. However, subtracting 1920 from the average of 1918 and 1919 gives –3.49°, –0.15°, and –0.55°C, respectively. Thus, MADI was about 3°C colder than the other two relative to the previous years. It appears that the retreating character of the index in the maximum thermometer tube began in 1920, rendering this year cooler than observed and so this summer was eliminated from MADI. Values from 1922 onward at MADI were in good agreement with other stations.

SCOT 1944. SCOT, one of the Far network stations, exhibited relatively warm temperatures in 1944, reporting the warmest anomaly in each summer month of 1944 relative to all 18 stations in the northern third of Alabama (including one that was even farther east and much higher elevation). SCOT was 2°C warmer than expected from comparisons with stations in the Near and Far datasets. The values for SCOT in 1944 were eliminated. There was no obvious explanation for this anomaly, but we are reminded that SCOT incurred more station moves and more incomplete months than any other examined here.

HSV1 1964. The 1964 JJA temperatures at the U.S. Weather Bureau office at the Huntsville airport show cooler than expected values by about 1.2°C vs the mean of 6 other nearby stations (see Fig. 3, arrow). August alone was about 2.0°C cooler than estimated from surrounding stations.

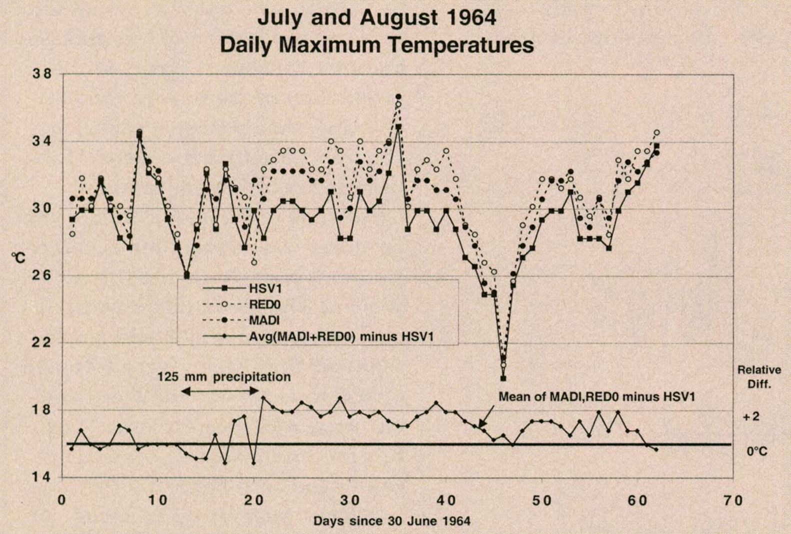

FIG. SB5. (top) Daily maximum temperatures for HSV1, RED0, and MADI for Jul and Aug 1964. (bottom) Difference between the mean of RED0 and MADI vs HSV1. Note the difference time series assumes its bias during a period of substantial rainfall.

To assess this situation further, given the fact the station was a fully staffed, 24-h office, I compared daily temperatures between HSV1 and two nearby stations, MADI and RED0. The time series of the differences between HSV1 and the average of the other two indicates a period in mid-July (12–20 Jul) in which HSV1 became relatively cool (Fig. SB5). During this period, approximately 125 mm of rain fell in the area. In 1964, HSV1 was located on the edge of a swampy area southwest of the town center. It seems likely that the evaporative cooling related to saturated or flooded ground in the immediate area prevented daily maximum temperatures at HSV1 from reaching values found at nearby sites that were characterized by better local drainage. (Note the somewhat positive correlation after 20 Jul between the absolute temperature and the difference time series.) By the end of Aug, the differences were random. Here is a case in which the moisture content of the surrounding land appears to have created a real reduction in monthly mean maximum temperatures. As a result, these monthly values at HSV1 in 1964 remain in the dataset and serve as an example of the random differences that affect and confound such reconstructions as attempted here.

ONEO 2000. JJA temperatures for ONEO were 2.37°C warmer in 2000 than in 1999. Seven stations reporting for both years indicated that 2000 was different by magnitudes that ranged from –0.09° to +0.78°C. No obvious explanation of the relative warmth of ONEO could be determined, but the large difference for this coop station implied a significant change, yet to be explained, did occur. The year 2000 was designated as a break in the time series of unknown cause.

A FURTHER CHECK USING NOT-SO-PERFECT FAR STATIONS

To obtain a mostly independent check on these results a second set of stations was gathered for the purpose of constructing a proxy for north Alabama. Relative to the present NWS office in Huntsville, the Far stations are Scottsboro (SCOT, 70 km east), the average of St. Bernard (STBE, 50 km south) and Oneonta (ONEO, 85 km SSE), and the average of Florence (FLOR) and Muscle Shoals (MUSC, ~85 km west). The averaging of the western stations (FLOR and MUSC) and southern stations (STBE and ONEO) is designed to allow equal weight to the data from three surrounding areas. In this way the data from SCOT to the east of the central core is equally weighted with data from the south and west. Data records began in 1894 (SCOT), 1896 (ONEO), 1908 (STBE), 1893 (FLOR), and 1941 (MUSC).

Scientists at NCDC selected stations for special designation based on objective criteria (i.e., length of record, limited relocations) to be included in the U.S. Historical Climatology Network (USHCN; Easterling et al. 1996.) It is interesting that none of the 13 Near stations was selected to be included in the USHCN. However, 3 Far stations qualified for USHCN (MUSC, STBE, and SCOT) though they have some significant shortcomings relative to our purpose here.

MUSC begins in 1941 and is thus relatively short for the immediate purpose. STBE begins relatively late (1908) and has a curious warming trend in the last few decades relative to all other stations. The station was located at a monastery that eventually became a college and thus has experienced localized development. Finally SCOT was initially broken into 27 identifiable segments, by far the largest number of breaks found among all of the stations examined. For a long period, the station was assigned to a radio broadcaster who moved numerous times and for which weekend observations in the summer were evidently difficult to obtain (thus many missings). I eventually reduced the number of segments to 14, assessing the changes individually to assure that little impact was evident by this reduction.

There is less confidence in this Far reconstruction due to the many changes (41 segments for 5 stations) and lack of data. In addition, these stations are separated by up to 155 km, and will be subject to real differences across these distances. Finally, with only five stations included in the time series, segment biases are less well characterized. For the pre-1940 period, the standard deviations of JJA averages about the calculated mean are 0.50°C for Far and 0.39° for Near. For post-1940 data, the values are 0.32° and 0.27°C for Far and Near, respectively. Note that the standard deviations of the Far data are determined from 3 values each year, 2 of which are already averages of 2 stations each (MUSC and FLOR, ONEO and STBE). Had the individual values of the 5 Far stations been used as in the Near comparison, the standard deviations would be even greater for Far.

The largest difference between the Near and Far time series is the relative warmth in Near for the period from 1920 to 1940 when only 3 of the 5 Far stations reported regularly (Fig. SB6). Outside of this period, the decadal differences are relatively small. Because of these relatively warm temperatures in the Near stations, the trend is more negative (–0.13°C decade^{-1}) than that of the Far stations (–0.08°C decade^{-1}).

There is little ability to assess the confidence of the trend estimate. I am encouraged, however, that the differences in 108-yr trends of the 13-, 14-, and 15-station networks in Near, having 2, 3, and 4 stations, respectively, in the 1893–1940 period, varied by less than 0.005°C decade^{-1}. However, there may be a set of peculiar coincidences in the production of b from the segment bias matrix that conspire to create an erroneous trend. I can only state that the production of b is unbiased and that I assumed errors in the production of the b were random and thus should largely cancel.

I also note that the trend of the time series of the adjusted Far network is more negative than that produced when assuming each Far station is a single segment (not shown.) The trend differences between the multiple- and single-segment time series were 0.06 and 0.10°C decade^{-1} for the Near and Far networks, respectively. Thus this method indicates consistent results for both Near and Far networks, suggesting spurious warming for JJA maximum temperatures resides in the unadjusted, mostly rural stations in north Alabama.

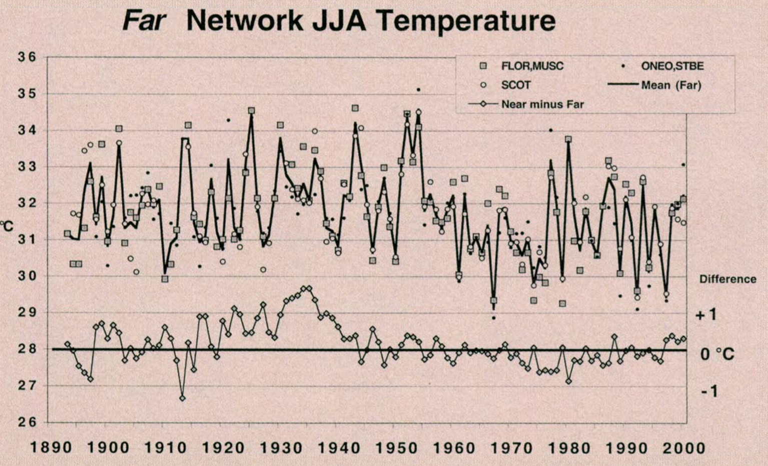

FIG. SB6. (top) Time series of JJA maximum temperatures for the Far network (line). Each symbol represents one of the three components of the average. (bottom) Difference (Near minus Far).

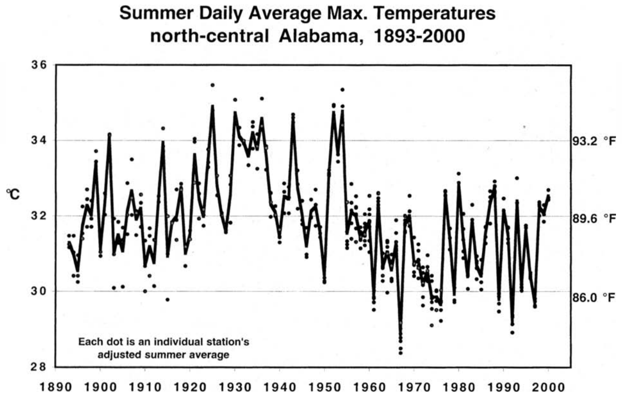

FIG. 4. Time series of the grand average of the Near + FLOR 14-station network. Individual station values are shown as filled circles (after bias removed relative to the NWS office in Huntsville).

coolest summer was probably 1967 (28.86°C), with 1992 (29.16°C) second coolest. The hottest single month was very likely July 1930 at 36.87°C since it was much warmer than the second and third warmest, 35.97°C (July 1952) and 35.87°C (August 1954). The trend since 1893 is –0.13°C decade^{-1}.[3]

I would point out that none of the nearby stations operating in 1925 are in operation today, and even if they were, many discontinuities have crept into the record as to raise suspicion. So, determining the year of the hottest summer is not a trivial matter and this is my best estimate.

By the time this analysis method is described to answer the question posed, the inquisitor is generally confused and interest is lost. The context of life in which many of our questioners live is characterized by absolute "winners" and "losers" (e.g., which year won the highest temperature contest?). We tend to disappoint them with our inability to provide definitive answers.

Climate data records require estimated levels of uncertainty (National Research Council 1999, 2000). This single idea is manifestly difficult to transfer to the public and the media who believe "uncertain" often means "erroneous" or "useless." One would never expect a headline writer to announce a story with, "Summer's Temps 7th to 12th Warmest since 1893 with 95% Confidence!" though it may be highly accurate. The implication of this study is that we often cannot provide an absolute answer to the media and the public with the confidence they typically expect from scientists who deal with observations from scientific instruments.

Finally, the reader is advised to be wary of pronouncements about extreme events, especially in localized areas (Pielke et al. 2000). The context in which the extreme is claimed must be rigorously assessed so that some level of confidence may be assigned to the event relative to a time period of at least a century. In perhaps more cases than we would admit, our ability to report an unambiguous extreme value is rather limited as shown by this simple example.

Acknowledgments. This work was partially supported by the U.S. Department of Transportation (DTFH61-99-X-00040). In addition, the author's wife, Babs Christy, graciously endured many evenings and weekends in which her dining room was cluttered with original observational forms from the stations studied herein. The personnel of NCDC were especially helpful, among those being Thomas Ross, Thomas Peterson, Thomas Karl, Richard Heim and David Easterling.

REFERENCES

Easterling, D. R., and T. C. Peterson, 1995: A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol.*, **15,** 369–377.

——, T. R. Karl, E. H. Mason, P. Y. Hughes, D. P. Bowman, R. C. Daniels, and T. A. Boden, 1996: U.S. Historical Climatology Network Monthly Temperature and Precipitation Data, Environmental Sciences Division Publ. 4500, ORNL/CDIAC-87, 263 pp. [Available from Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN 37831.]

Hansen, J., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, and T. Karl, 2001: A closer look at United States and global surface temperature change. *J. Geophys Res.*, **106,** 23 947–23 963.

Karl, T. R., C. N. Williams, Jr., P. J. Young, and W. M. Wendland, 1986: A model to estimate the time of observation bias associated with monthly mean maximum, minimum and mean temperatures for the United States. *J. Climate Appl. Meteor.*, **25,** 145–160.

——, H. F. Diaz, and G. Kukla, 1988: Urbanization: Its detection and effect in the United States climate record. *J. Climate,* **1,** 1099–1123.

——, C. N. Williams, F. T. Quinlan, and T. A. Boden, 1990: U.S. Historical Climatology Network Serial Temperature and Precipitation Data, Environmental Sciences Division Publ. 3404, ORNL/CDIAC-30,

---

[3] A check of NCDC's Climate Division temperature data of daily means (i.e., average of maximum and minimum) shows that since 1895 the warmest summers were 1952 (27.89°C), 1943 (27.70°C), 1954 (27.46°C), and 1934 and 1936 (27.24°C). The trend was –0.09 C decade^{-1}.

83 pp. [Available from Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN 37831.]

National Research Council, 1999: *Adequacy of Climate Observing Systems.* National Academy Press, 51 pp.

——, 2000: *Issues in the Integration of Research and Operational Satellite Systems for Climate Research: Implementation.* National Academy Press, 82 pp.

Peterson, T. C., T. R. Karl, P. F. Jamason, R. Knight, and D. R. Easterling, 1998a: First difference method: Maximizing station density for the calculation of long-term global temperature change. *J. Geophys. Res.,* **103,** 25 967–25 974.

——, and Coauthors, 1998b: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.,* **18,** 1493–1517.

——, K. P. Gallo, J. Lawrimore, T. W. Owen, A. Huang, and D. A. McKittrick, 1999: Global rural temperature trends. *Geophys. Res. Lett.,* **26,** 329–332.

Pielke, R. A., Sr., T. Stohlgren, W. Parton, N. Doesken, J. Money, L. Schell, and K. Redmond, 2000: Spacial representativeness of temperature measurements from a single site. *Bull. Amer. Meteor. Soc.,* **81,** 826–830.

Quayle, R. G., D. R. Easterling, T. R. Karl, and P. Y. Hughes, 1991: Effects of recent thermometer changes in the cooperative station network. *Bull. Amer. Meteor. Soc.,* **72,** 1718–1723.

With the development of meteorological science and the continual refinement of the technologies used in its practical application, the need to produce a new edition of the *International Meteorological Vocabulary* (IMV) became evident (the original edition was published in 1966). This volume is made up of a multilingual list of over 3500 terms arranged in English alphabetical order, accompanied by definitions in each of the languages (English, French, Russian, and Spanish) and an index for each language. This new edition has been augmented with numerous concepts relating to new meteorological knowledge, techniques, and concerns. It should help to standardize the terminology used in this field, facilitate communication between specialists speaking different languages, and aid translators in their work.

WMO No. 182, 784 pp., softbound, color-coded index, $95 (including postage and handling). Please send prepaid orders to: WMO Publications Center, American Meteorological Society, 45 Beacon St., Boston, MA 02108-3693. (Orders from U.S. and Canada only.)

World Meteorological Organization

International Meteorological Vocabulary

Downloaded from http://journals.ametsoc.org/bams/article-pdf/83/5/723/3733428/1520-0477(2002)083_0723_wwths_2_3_co_2.pdf by guest on 10 October 2020