Urbanization: Its Detection and Effect in the United States Climate Record

THOMAS R. KARL
National Climatic Data Center, NOAA, Asheville, North Carolina

HENRY F. DIAZ
Environmental Research Laboratories, NOAA, Boulder, Colorado

GEORGE KUKLA
Lamont-Doherty Geological Observatory, Columbia University, Palisades, New York

(Manuscript received 18 January 1988, in final form 29 June 1988)

ABSTRACT

Several equations were developed that related the effect of urban growth, measured by increasing population, to the mean seasonal and annual temperature: the diurnal maximum, minimum, average, and range. These equations were derived from a network of 1219 stations across the United States, which were analyzed for the years 1901–84. The results indicate that urban effects on temperature are detectable even for small towns with populations under 10 000. Stations with populations near 10 000 are shown to average 0.1°C warmer for the mean annual temperature than nearby stations located in rural areas with populations less than 2000. Urbanization decreases the daily maxima in all seasons except winter and the temperature range in all seasons. It increases the diurnal minima and the daily means in all seasons.

The equations indicate that, for the annual mean temperature, urbanization during the twentieth century accounts for a warm bias of about 0.06°C in the U.S. Historical Climatology Network (HCN). Due to the large number of stations located in sparsely populated areas [(over 85% (70%) of all stations had a 1980 population of less than 25 000 (10 000)], the impact of urbanization is not large in relation to decadal changes of temperature in the United States. The average heat island impact during the period 1901–84 for the HCN is largest for the daily minima (0.13°C) and the temperature range (−0.14°C), while the impact on the daily maxima (−0.01°C) is an order of magnitude smaller.

1. Introduction

The instrumental records of monthly, seasonal, and annual averages of land surface temperatures are useful in a wide range of applications, from studies of important seasonal and multiyear climate anomalies and fluctuations (Ropelewski and Halper 1986; Yarnal and Diaz 1986; Karl and Young 1987; Karl 1988), to assessments of long-term (nineteenth and twentieth century) temperature trends (Diaz 1986; Jones et al. 1986; Ellsaesser et al. 1986; Karl et al. 1984). The latter is an important and essential aspect of efforts aimed at the early detection of the greenhouse effect (Wigley et al. 1985). Considering the importance and usefulness of such data, it is surprising to find a dearth of studies aimed at the identification and quantification of the effects of urbanization, not only for large urban areas but for cities of all sizes, especially since urbanization and the greenhouse effect are expected to produce similar signals in the climate record.

A considerable number of studies (Oke 1973; Landsberg 1975; Oke 1979; Landsberg 1981; Oke 1982; WMO 1986) have addressed urban heat as a physical phenomenon and quantified various aspects of maximum instantaneous urban-to-rural temperature differences. Unfortunately, these results are difficult to link to seasonal and annual mean temperatures. Other studies have attempted to quantify the effect on the temperature record of increasing levels of urbanization in major cities (Mitchell 1953; Dettwiller 1970; Fukui 1970; Cayan and Douglas 1984; Duchon 1986; Goodridge 1985; Kukla et al. 1986; Chow 1986). These studies demonstrate the importance of isolating and removing from the climate record, for some applications (i.e., climate change detection, seasonal and monthly forecasts, etc.), the anomalous warmth produced by major cities. At the same time, the question remains regarding the level of urbanization which can be tolerated in the construction of a temperature record with an acceptably low bias for climate change studies.

There are many reasons why this important issue has not been adequately addressed in the construction of large-area averages of temperature. The difficulty in quantifying "urbanization" has no doubt frustrated

Corresponding author address: Thomas R. Karl, Climatological Analysis Division; NOAA/NCDC, Federal Building, Asheville, NC 28801.

many investigators. In addition, a more physically oriented approach necessitates the study of the urban climate of specific cities, each with its own unique characteristics. As pointed out by Oke (1979), the internal pattern of the heat island of each city is often dominated by microscale features related to land use and building density. In large-scale studies of climate change using hundreds or even thousands of stations, the microclimate of each station cannot, at present, be properly parameterized. This presents a major obstacle to the quantification of the heat islands produced by small cities and necessitates the use of population as a measure of urbanization. The noise introduced by the microclimate of each city, however, may disguise the effect of urbanization.

One method of trying to overcome such difficulties is to use very large samples, i.e., many pairs of small cities and rural stations. Until recently, such data were not readily accessible. Fortunately, the Historical Climatology Network (HCN) provides a relatively dense network of long-term stations (1219) across the contiguous United States (Quinlan et al. 1987). Many stations in this network are located in small cities in remote areas [over 85% (70%) of all stations had a 1980 population of less than 25 000 (10 000)]. Data from this network provide an opportunity to assess the effect of urbanization on the climate record for urban areas of various sizes.

2. Data

The HCN contains monthly averages of maximum, minimum, and average temperature for 1219 stations with varying periods of record, most beginning either in the early 1900s or the late 1800s. In the United States and in many other countries, the daily average temperature is defined as the sum of the maximum and minimum divided by 2. At the time of this study, these data were available through 1984. The data have been subjected to rigorous quality control procedures and to various other procedures that adjust for *documented* potential discontinuities (station moves, new observation schedules, new instruments, etc.). Karl and Williams (1987) and Karl et al. (1986a) discuss the methodology used to adjust the data for these discontinuities. Karl and Williams (1987) use the method of differences between neighboring stations and Monte Carlo simulations to assess the significance of any potential discontinuity, and Karl et al. (1986) provide a model for eliminating the bias associated with varying observation schedules at cooperative stations. Quinlan et al. (1987) provide information regarding the quality control procedures used in making the data computer compatible as well as an objective summary of the integrity of each station's record for a variety of factors. They rank the stations in the HCN in terms of the statistical confidence interval of the adjustments they apply (or do not apply) for potential discontinuities and the percent of adjusted data that had to be estimated using the Karl and Williams (1987) methodology. For some early periods, potential discontinuities are either too frequent or nearby stations too distant for any adjustments to be reliably attempted. These data were not used. Relationships developed in this study for the *detection* of the urban heat island are derived from HCN stations that were ranked in the upper 90% with respect to the average width of the confidence intervals for adjustments due to discontinuities. In addition, no station had more than 25% of its data replaced with estimates from nearby stations because of missing observations or frequent moves.

The magnitude of the heat island bias in the HCN was estimated using these stations as well as a network that included all stations, regardless of the confidence interval of the adjustments and the quantity of estimated data (see section 6).

3. Procedures

The purpose of this study is to produce a method by which the urbanization effects of even small cities in midlatitudes (in the United States, at least) can be removed prior to their use in forming regional- or continental-scale spatial averages of temperature for analyzing secular climate change. The parameter chosen to represent urbanization is the population of the city or metropolitan area where the station is located. This is not the most desirable physical quantity for representing urbanization around the climate station, but it is one of the few documented statistics that is readily available for the past century and over much of the world. Differences in the thermal conductance between urban and rural surfaces may be important, but Oke (1981) argues that there is little data to support large differences. Indeed, Oke (1981) provides data to support the notion that the canyon sky-view factor (the fraction of the overlying hemisphere occupied by sky) in the city is the dominant mechanism that produces urban heat islands during nights with calm, cloudless skies, where anthropogenic heat is of negligible importance. Unfortunately, it is not possible to calculate this quantity from existing station histories. In the United States, however, the populations of every incorporated city has been defined by the U.S. Census Bureau since 1790. In this study, these populations were used as a measure of urbanization. When populations of metropolitan areas were defined by the U.S. Census Bureau, they were used to represent urbanization in contrast to the population within city limit boundaries.

A total of 941 stations were identified from the HCN as having complete records going back to 1941 and 541 with records going back to 1901 in which adjustments for potential discontinuities could be assessed. It must be emphasized that these stations underwent extensive scrutiny before they were used (Karl et al. 1986a; Karl and Williams 1987). From these stations a set of urban/rural station pairs was formed so that the distance between the urban and rural sites was no more than 100 km and each rural station was at least 30 km from the urban station. In this regard, we suspect that most of the rural stations were outside the urban plume. Indeed, as Lowry (1977) indicates, in the surface boundary layer the thermal effects of the city have a limited area of extent. The conceptual basis for this is described in Oke (1976, 1982) and Clarke (1969). During the evening and nighttime, the urban thermal plume quickly rises above the surrounding rural surface boundary layer.

We have arbitrarily defined rural stations as those with a 1980 population of less than 2000. Figures 1–4 depict these station pair groupings for the periods 1901–84 and 1941–84 for both seasonal mean temperature and the means of the maximum, minimum, and diurnal temperature range (the latter being the difference between the maximum and minimum temperature values). During the 1901–84 period these form a disproportionate share of the stations in the eastern United States, but for the 1941–84 period the distribution is more symmetric. In the HCN, adjustments for discontinuities were made independently for each element (maximum, minimum and mean), so that the number of pairs can vary because of the differing quality and availability of data for the different elements. Station pairs among the maximum, minimum, and temperature range, however, were not allowed to vary. Stations were used only if *both* the maximum and minimum passed the quality checks. As depicted in Figs. 1–4, a different set of stations was used for the maximum, minimum, and temperature range compared to the mean temperature.

The question of an appropriate method for segregating the urban effect from the background climate arises. Lowry (1977) provides a general framework describing the assumptions required when using, among other methods, urban-to-rural temperature differences as a measure of the urban heat island. These temperature differences have been used with two different techniques. First, urban-to-rural temperature differences were related to a measure of urbanization at the urban station. The rural station was assumed to be

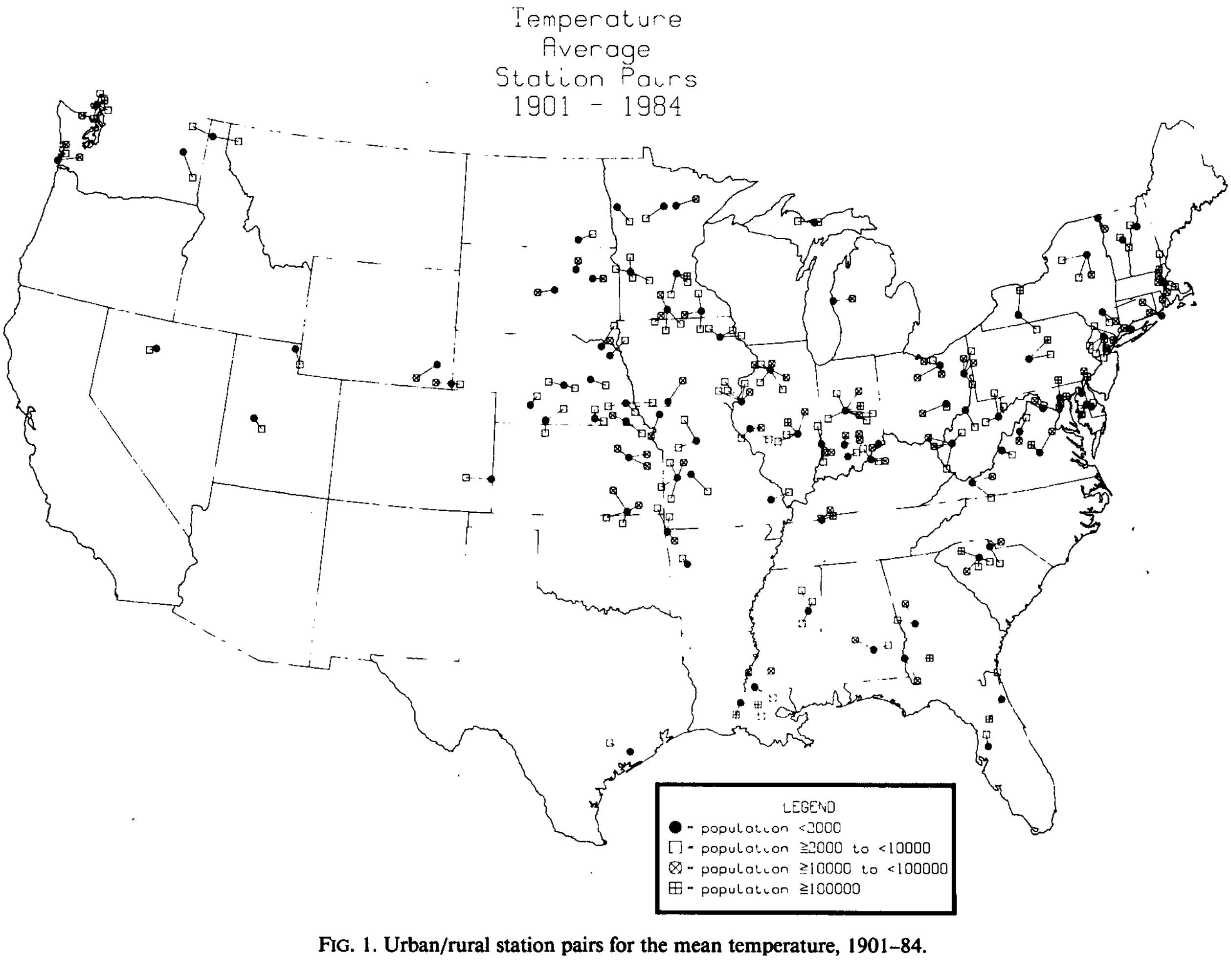

FIG. 1. Urban/rural station pairs for the mean temperature, 1901–84.

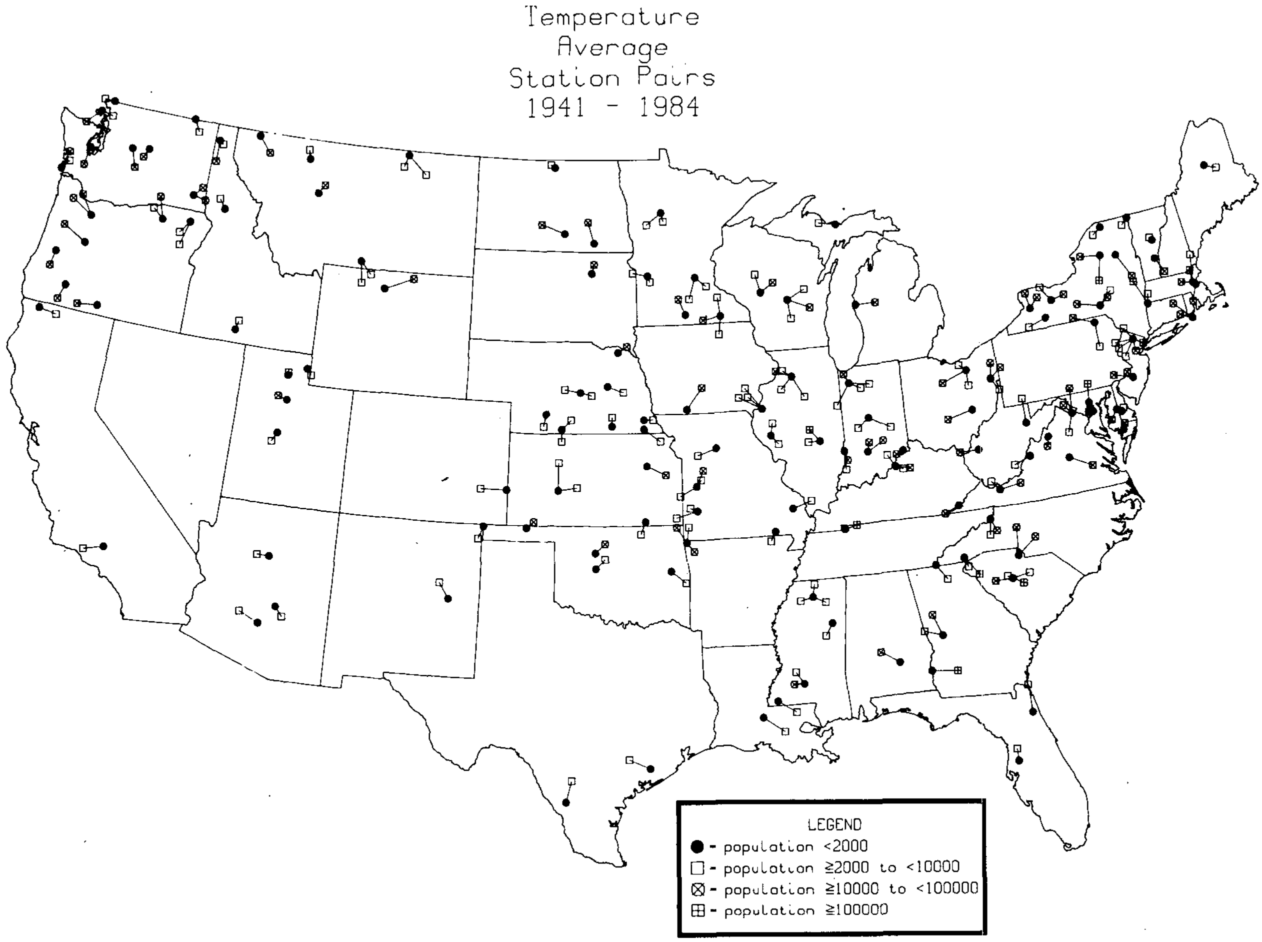

FIG. 2. As in Fig. 1, except for 1941–84.

uneffected by urbanization. In this approach, which we define as the basic Time Averaged Method (TAM), differences between the two stations in local landscape (topography, vegetation, water bodies, soil types, etc.) and background climate (latitude, continentality, etc.) must also be assumed to be unimportant. Clearly, this is a poor assumption, one we will need to address. In the second approach, which we define as the Time Rate of Change Method (TCM), the time trends of the urban-to-rural differences of temperature are related to a measure of urbanization at the urban station (cf. Mitchell 1953, 1961). In this approach, the differences in local landscape and background climate are minimized, but as Lowry (1977) points out, even here, one must assume that the probability of weather regimes or weather types remains constant over time. In addition, and perhaps in some instances this is a more severe assumption, the homogeneity (no significant changes of instruments, instrument location, vegetation, observing schedules, etc.) of the temperature records at the two sites must continue throughout the period of comparison. Since neither method of detecting the effects of urbanization is ideal, we used both, but with some significant variations in the basic TAM due to the assumptions of the background climate and landscape at each station.

After all urban-to-rural station pairs were identified, temperature differences were calculated, for each season and annually, for the urban minus the rural station temperature. To assess the effects of urbanization on the temperature record, these temperature differences were used with the two techniques. In the TCM, the changes in the urban-to-rural temperature differences were categorized over time. That is, time series of urban minus rural temperature differences were calculated and categorized as defined in Table 1. The rural stations chosen had a relatively low population throughout the period of record, while the urban stations were selected solely on the basis of their 1980 population. This means that the rate of growth of these urban stations is only limited by their 1980 population, but growth rates can vary from station to station within categories.[1] In ad-

[1] The effects of varying population growth rates within categories were also tested using this method.

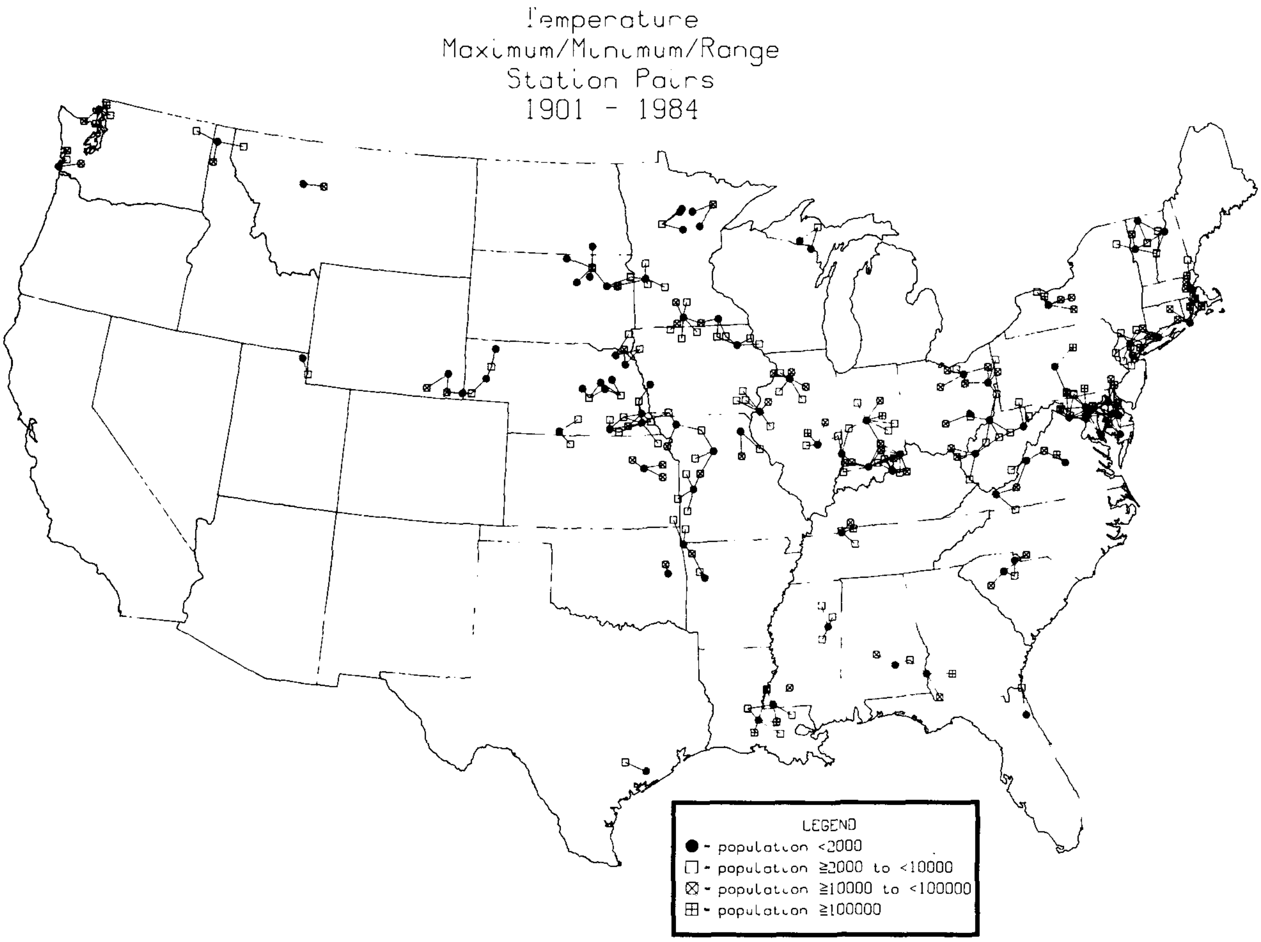

FIG. 3. As in Fig. 1, except for the maximum, minimum, and diurnal temperature range.

dition to the Mitchell (1961) study, many others have used this approach to identify the effect of urbanization in climate time series (Duchon 1986; Cayan and Douglas 1984; Kukla et al. 1986; Goodridge 1985). In the alternative to the basic TAM, hereafter referred to simply as TAM, averages of the urban-to-rural temperature difference over small time segments (several years) after removing the effects of elevation and latitudinal differences, were related to the population of the urban station. (As previously noted, the rural station's population is known to be less than 2000.)

Identical station pairs were used with both techniques for the 1941–84 period of record (the 1901–84 station pairs were not used in the TAM). Another separate set of data was included to make sure the adjustment procedure used by Karl and Williams (1986) did not mask heat island biases. The original observations (adjusted for time of observation biases) were used in the TCM for the station pairs depicted in Figs. 1–4.

In the TCM one must identify the time rate of change of the seasonal and annual means of the urban-to-rural temperature differences. We accomplished this by differencing period means. The 1901–42 mean difference was subtracted from the 1943–84 mean; likewise, the mean for the period 1941–62 was subtracted from the mean for 1963–84. Other methods, such as ordinary least square trends and resistant (robust) least square trends (Emerson and Hoaglin 1983), were considered but dismissed because of the nature of the time series derived using the TCM. Irregular fluctuations were often deeply embedded in the time series, exaggerating the leverage of the time series near the beginning and end points associated with trend analysis (although resistant methods would tend to minimize this to some extent). Because of the relatively large number of pairs in each category of population (Table 1), it was possible to assess the statistical significance of this difference across the various time periods by using a paired difference t-test with a null hypothesis, "The difference of the mean of all stations aggregated within a category across the two time periods is zero," versus an alternate hypothesis, "The difference of the mean across the two time periods is positive." In the paired difference t-test, the standard deviation is calculated on the basis of the station-to-station (pairs) variability of the urban-to-ru-

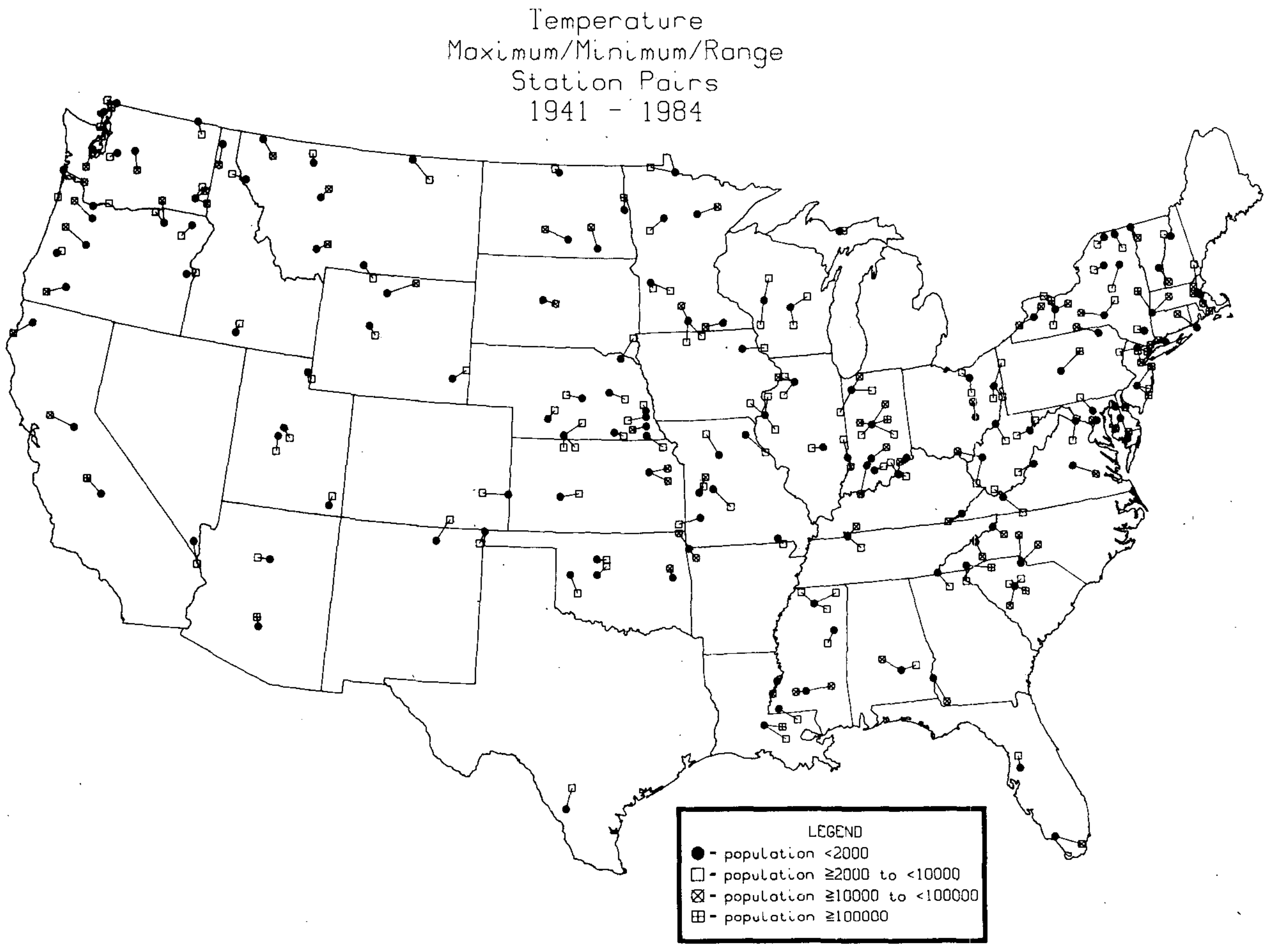

FIG. 4. As in Fig. 3, except for 1941–84.

ral temperature difference between the two time periods. In this method, factors such as the differences between urban and rural station pairs, i.e., topography, latitude, continentality, etc., are not very important because *only* the changes within each station pair are considered, not the absolute difference.

The TAM avoids the disadvantages associated with the use of long time series to assess the effects of urbanization. Instead, time averages of a small group of years are used to calculate mean urban-to-rural temperature differences. Specifically, the *mean* urban-to-rural temperature difference ($\bar{T}_{u-r}$) was calculated separately for 9 years centered around both 1950 and 1980, i.e., 1946–54 and 1976–84, respectively. The differences were then related to the urban population. This approach minimizes the effects of undocumented

TABLE 1. The mean 1980 population (left of solidus) and the number of urban-to-rural station pairs which includes the total for all comparisons under the <2000 category (right of solidus) is given for one rural and three urban categories. Abbreviations are NA, not applicable; MAX, maximum temperature; MIN, minimum temperature; RNG, temperature range (maximum − minimum); AVERAGE, (MAX + MIN)/2.

		1980 Population			
Period of record	Variable	Rural <2000	Small urban 2000–9999	Medium urban 10 000–100 000	Large urban >100 000
1941–84	Average	750/414	5030/233	23 362/150	740 403/31
1941–84	Max/Min/Rng	724/407	5084/226	23 365/148	805 202/32
1901–84	Average	840/213	5086/120	24 416/75	987 474/18
1901–84	Max/Min/Rng	841/305	5250/152	25 652/128	1 296 767/25

changes to the stations but, as previously mentioned, it introduces other related nonurban sources of variability between pairs, i.e., differences in elevation, latitude, continentality, etc. To the extent possible, this additional variability was removed from the urban-to-rural temperature differences by modifying the basic TAM. This was accomplished by using the observed-predicted residuals ($\bar{T}^*_{u-r}$) of $\bar{T}_{u-r}$ from multiple regression equations, in which the differences in latitude (part of the background climate) and elevation (part of the local landscape climate) of the station pairs were used as independent variables and the $\bar{T}_{u-r}$ as the dependent variable. These equations were developed on a seasonal and annual basis over the entire United States. Such an approach cannot address all the important local landscape climate characteristics, such as a river valley location versus a plateau, a mountain slope versus a mountain valley, a sea breeze versus no sea breeze, etc., but it does help to minimize two of the major differences in climate among the urban-to-rural station pairs.

Throughout this study we average across many station pairs, which mitigates much of the noise introduced by local landscape climate variability and background climate if this variability occurs in a near-random manner across the station pairs. We have no reason to believe that it does not.

4. Results

a. *Detection of urbanization using the Time Rate of Change Method (TCM)*

The year-to-year, urban-to-rural temperature differences ($\bar{T}_{u-r}$) for the 1901–84 time period and for each of the four elements—annual average, maximum and minimum temperature, and temperature range—are depicted in Figs. 5–8 for various categories of 1980 urban population. In Fig. 5, it is apparent that there is considerable year-to-year variability of $\bar{T}_{u-r}$ for all categories, and that there are also significant decadal or multidecadal fluctuations. For the lowest category of urban population, 2000–10 000, it is difficult to detect a systemic change in $\bar{T}_{u-r}$, but for other categories there is a general tendency toward larger positive $\bar{T}_{u-r}$ in the later years of record. It must be emphasized that these time series are compiled from many station pairs and that some time series of individual urban-to-rural station pairs look quite different from those presented here. Although it can be logically argued that abrupt shifts in the times series of urban–rural temperatures are characteristic of the steplike development and architectural spurts common in large cities, it is difficult to explain the fluctuating or vacillating characteristics of the time series using this reasoning. It is conceivable that these are part of natural climate fluctuations (changes in the frequency of weather types) as envisioned by Lowry (1977). This is especially pertinent considering that no time series presented in Fig. 5 consists of fewer than 18 pairs and usually includes many more.

The maximum temperature (Fig. 6) for all categories with a 1980 population of less than 100 000 displays no tendency toward increasing $\bar{T}_{u-r}$ and in fact, the category with a population of 10 000 to 25 000 exhibits a tendency toward decreasing urban–rural temperature differences over time. The category with 100 000 or more shows an abrupt rise during the mid-1950s, with a tendency in later years to return to differences of magnitude prior to the 1950s. When differences between the two 42-yr periods (1901–42 versus 1943–84) and the two 22-yr periods (1941–62 versus 1963–84) are calculated, larger positive values of $\bar{T}_{u-r}$ can be found in the latter period. The nature of the abrupt rise in the 1950s and the fall in the 1980s, however, suggests that this difference may be due to causes unrelated to urbanization effects.

All categories of population exhibit a general increase of $\bar{T}_{u-r}$ with respect to the minimum temperature (Fig. 7). This is apparent for the categories with a 1980 population of 2000–10 000, 10 000–25 000, and >100 000 categories. The trend in the 10 000–25 000 category, however, may simply be an overestimate of the urban effect, since temperatures are shown to be lower in urban areas than in rural areas prior to 1925, and there is no trend for the 25 000 to 100 000 category until the mid-1960s. As stated earlier, systematic differences in local landscape and background climatic differences are not considered in this method, so it is conceivable that urban areas could have cooler temperatures than nonurban areas, although, in addition, multiyear fluctuations of the frequency of weather types affecting these differences must also be considered. Nevertheless, the general increase in the mean $\bar{T}_{u-r}$ is seen primarily as a consequence of changes in minimum (or nighttime) temperatures.

The effects on the diurnal temperature range reflect the nonuniformity of the trends of $\bar{T}_{u-r}$ for the maximum and minimum (Fig. 8). There is a definite trend toward a decreasing temperature range in the 10 000 to 25 000 category. Here again, a large portion of the decrease in this category is due to the rise in the minimum (Fig. 7), which may be exaggerated. This is also true in the 2000 to 10 000 category, but only for the period 1901–40. The decrease in range for other categories in less apparent.

Figures 9 and 10 summarize the magnitude of $\bar{T}_{u-r}$ for each variable and category identified in Table 1 with respect to the differences of $\bar{T}_{u-r}$ for the first half and second half (averages) of the 1901–84 and 1941–84 time periods. The most important result that can be derived from these figures in terms of the TCM is the fact that the *average* temperature appears to be slightly affected by urbanization even down to the lowest category, 2000 to 10 000; these results, however, are not statistically significant until the next category, 10 000 to 100 000, is considered. For the *minimum*

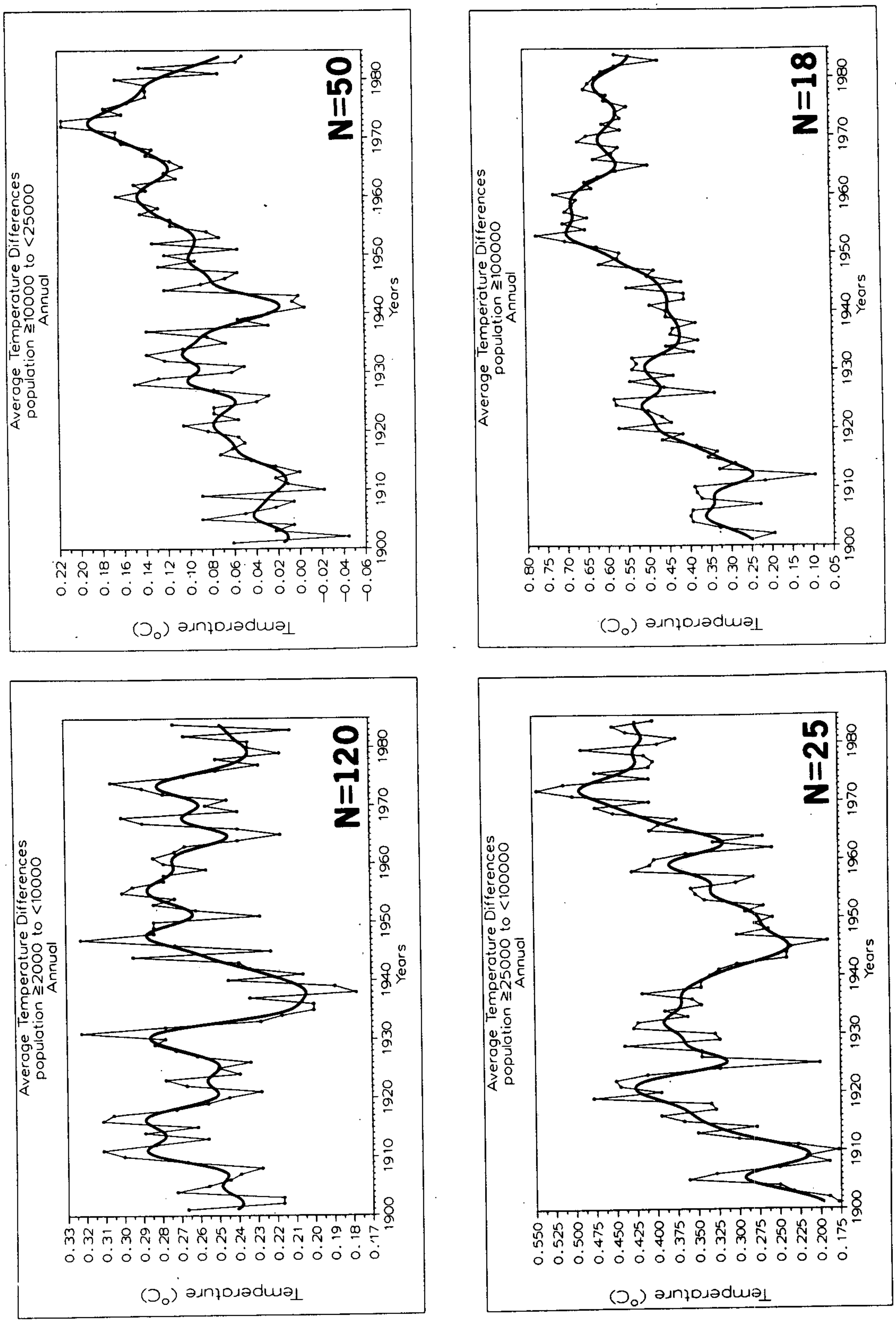

FIG. 5. Time series of the average urban minus rural (population < 2000) temperature difference for various categories of 1980 population at the urban station. Smoothed curve is a nine-point binomial filter with double weighted endpoints. Here N is the number of station pairs used in the graph.

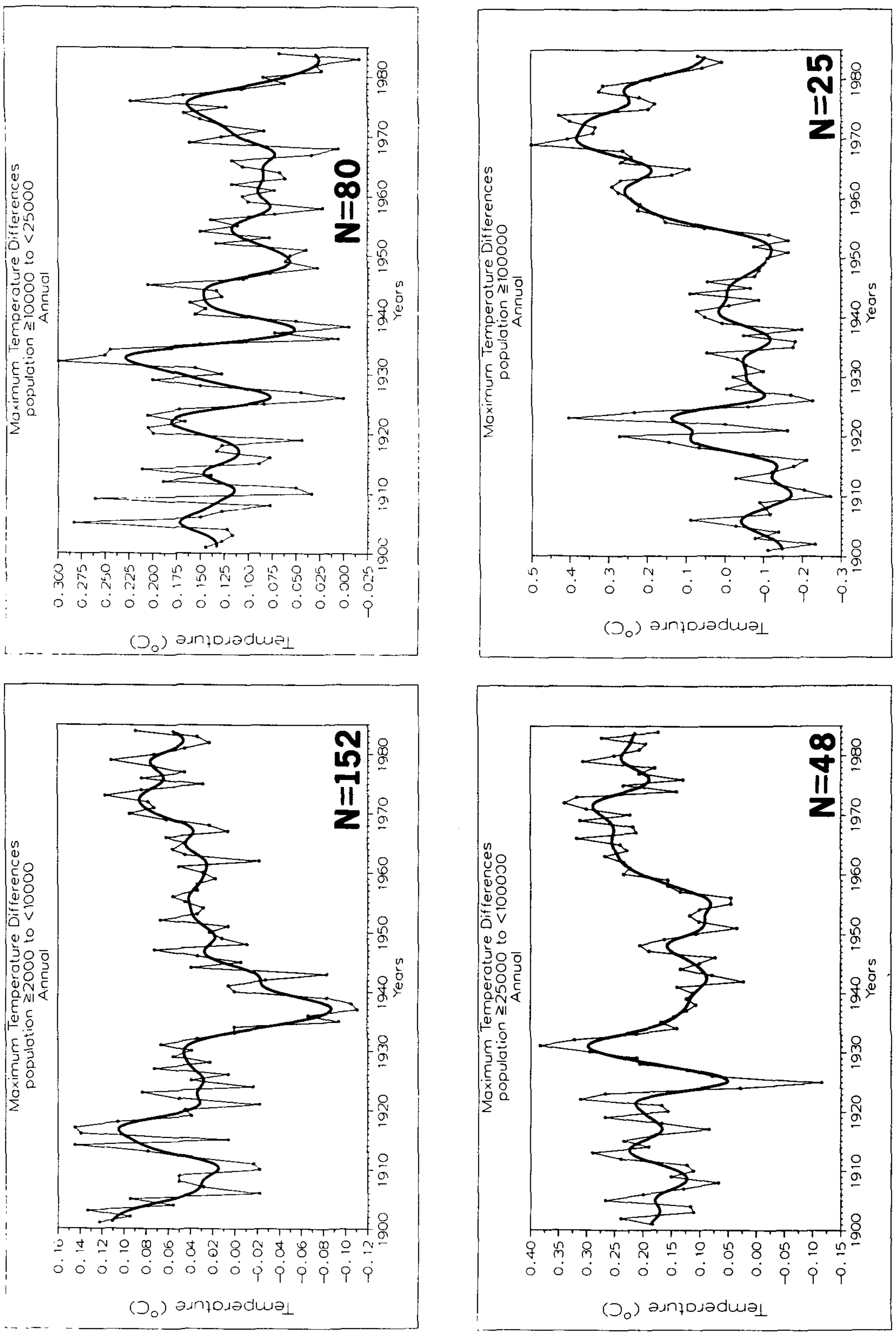

FIG. 6. As in Fig. 5, except for the mean maximum.

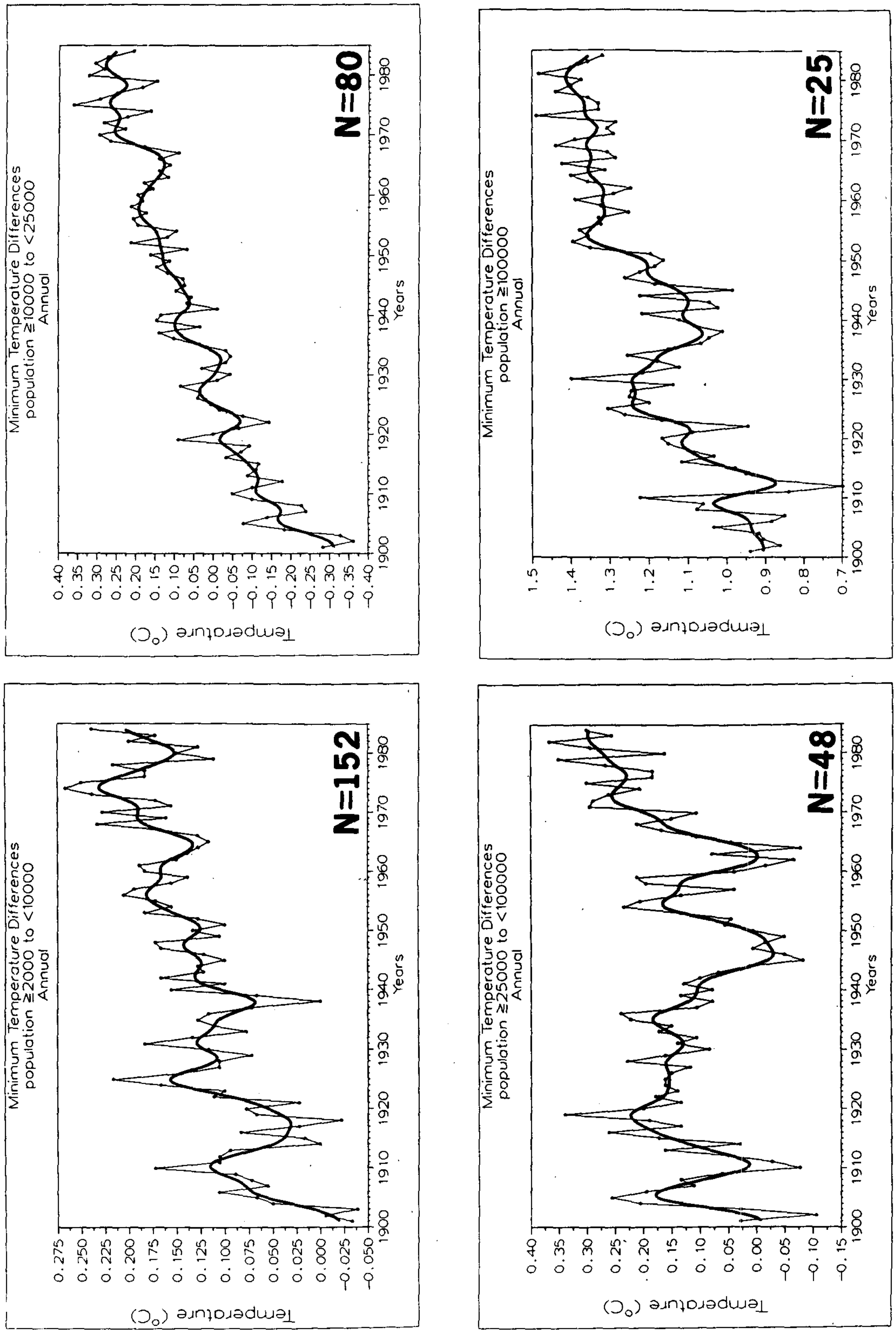

FIG. 7. As in Fig. 5, except for the mean minimum.

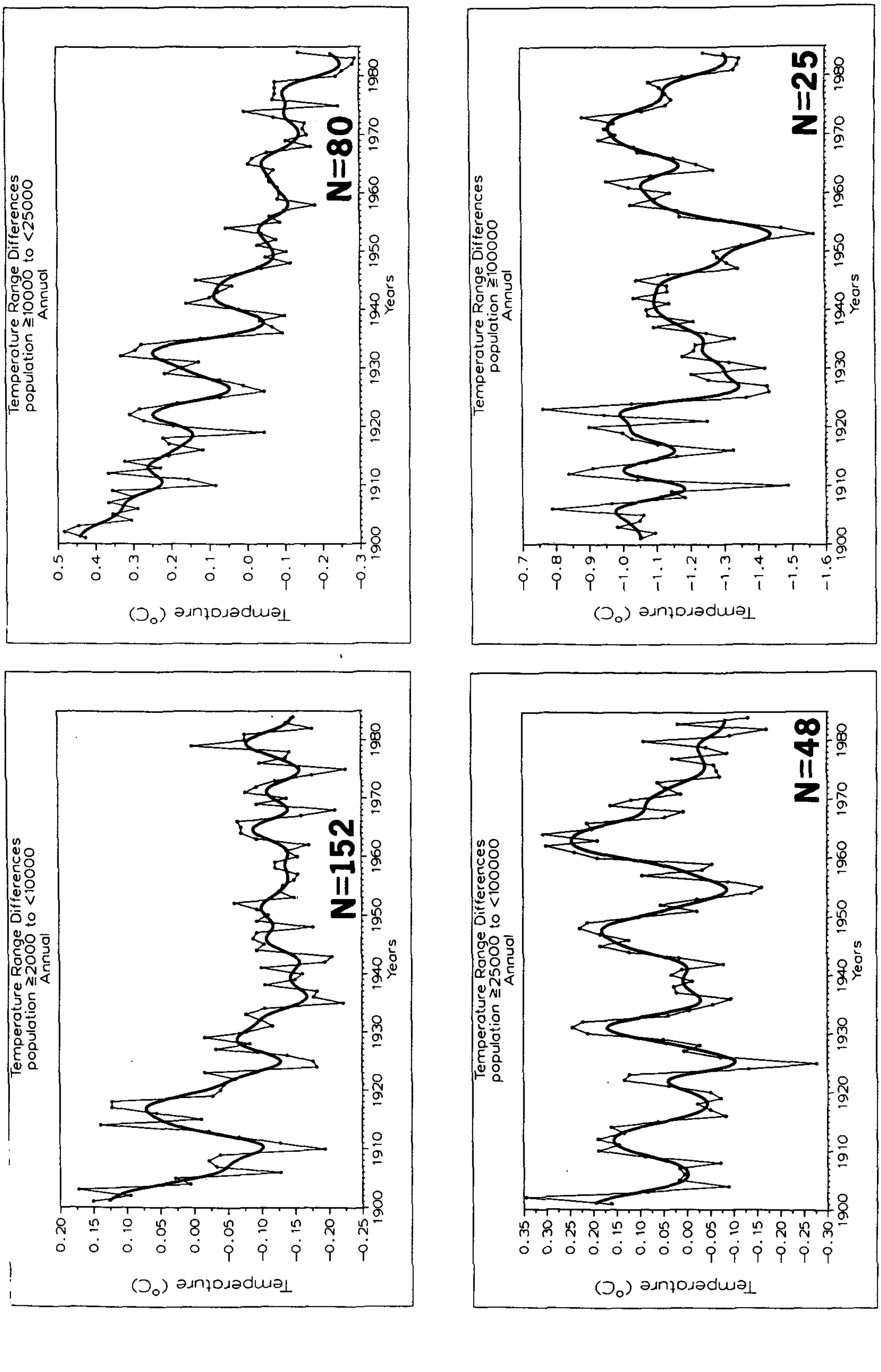

FIG. 8. As in Fig. 5, except for the mean diurnal range.

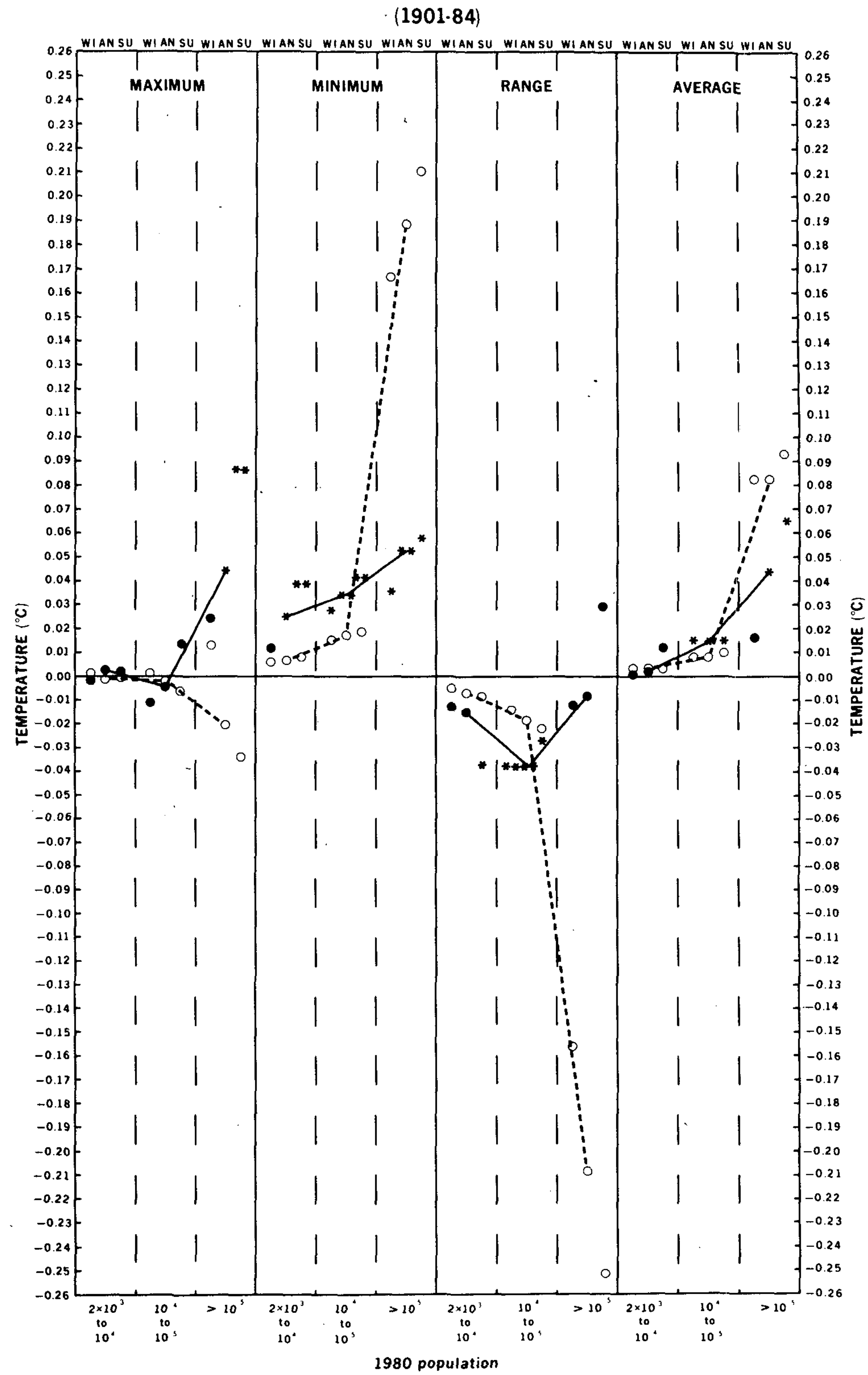

FIG. 9. Relative change in temperature (°C/decade) at urban stations compared to rural stations during the years 1901–84. Here WI implies winter; AN, annual; and SU, summer. Open (blackened) dots and dashed (solid) lines pertain to the TAM (TCM). Dashed (solid) lines connect annual values for the TAM (TCM) for the three population categories. For the TCM blackened dots are replaced by an asterisk when significant at the 10% level and double asterisks at the 1% level.

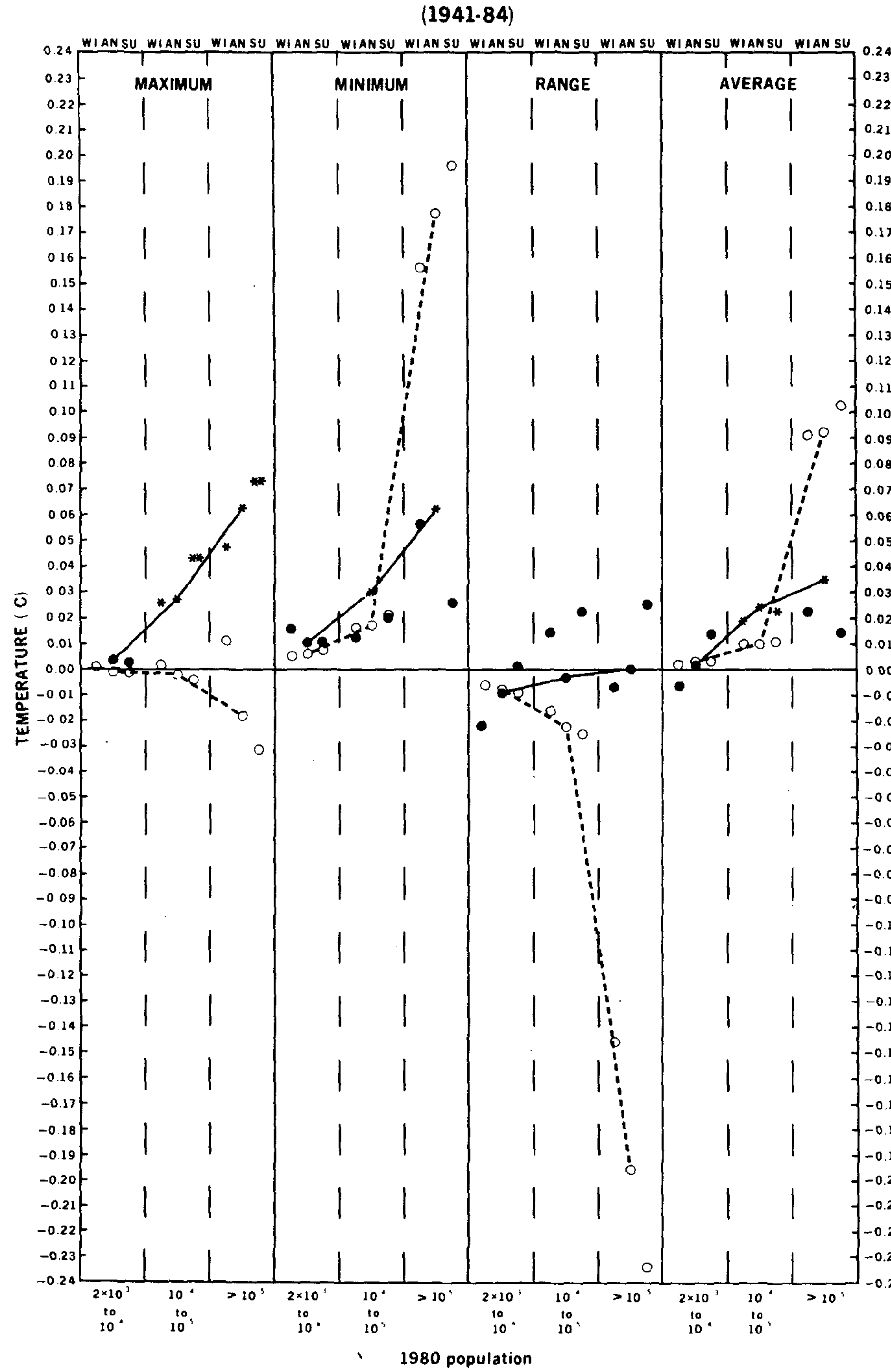

FIG. 10. As in Fig. 9, except for the 1941–84 time period.

temperature the increase in the urban–rural temperature difference can be found to be statistically significant even down to the 2000 to 10 000 population category. The inconsistency of the results for the *maximum* temperature for the different population groups is puzzling, especially for the 1941–84 time period. Previous work, summarized by Landsberg (1981), suggests that the maximum temperature is substantially less affected by urbanization than the minimum. Figures 5–8 indicate that a significant portion of the change of $\bar{T}_{u-r}$

is in the form of fluctuations or vacillations, as opposed to gradual increases that would reflect general increases in population or urbanization aggregated across many stations. Noting the uncertainty of the maximum temperature, the *temperature range* seems to have been only slightly affected by urbanization during the 1941–84 period but generally decreased in an erratic manner with respect to the urban stations over the 1901–84 period.

In order to determine if there were important regional differences with respect to the TCM, the analysis procedure was repeated by dividing the country in half, a northern network and a southern network, using the 40° parallel as a boundary. No significant differences could be detected. Similarly, it was divided into eastern and western networks, using the 90° meridian as a boundary, but again, no significant differences were found.

When the same analysis was repeated with original observations, uncorrected for potential discontinuities but corrected for time of observation biases, the relative warming of the stations located in urbanized areas was found to be approximately the same or smaller than that obtained from the adjusted data. This implies that the procedure used by Karl and Williams (1987) to adjust for potential discontinuities did not appreciably mask or smear urban warming effects across the network.

The rate of population change did not help explain the low-frequency fluctuations in the time series. The rate of change was used to categorize the subsets based on the population from 1920, 1950 and 1980. The logarithms and ratios of the changes were compared to the temperature difference series but no significant relationship was found. Many other transformations are possible, but only these two were tested.

b. *Detection of urbanization using the Time Average Method (TAM)*

1) Development of regression equations

Prior to the development of the relationship between instantaneous heat island intensity and canyon sky-view factors as a measure of urbanization, Oke (1976) and other workers used empirical methods to develop relationships between city size and the instantaneous maximum $\bar{T}_{u-r}$. The goal in this part of the analysis is to produce similar relationships, but for seasonal and annual differences of temperature. Once developed, these relationships can be used to remove the effects of increasing population (urbanization) in large-scale studies of climate change. For this reason, we developed regression coefficients for various regression equations. They are given in Table 2. Each equation was developed separately for both 1950 and 1980, and subsequently, by combining data for both 1950 and 1980[2] after some of the effects of local landscape and background climate (elevation and latitude) were considered. In the process of formulating these equations, several problems emerged.

TABLE 2. Coefficient a and the correlation (R) of regression equations using all populations of the form: $\bar{T}^*_{u-r} = a$ (urban population)$^{0.45}$, where $\bar{T}^*_{u-r}$ is the urban–rural temperature difference in °C.

Season	Coefficient a	Coefficient 95% confidence interval: Upper	Lower	R
		Average		
Winter	1.81×10^{-3}	2.01×10^{-3}	1.61×10^{-3}	0.53
Spring	1.47×10^{-3}	1.64×10^{-3}	1.30×10^{-3}	0.50
Summer	2.03×10^{-3}	2.35×10^{-3}	1.71×10^{-3}	0.40
Fall	2.01×10^{-3}	2.21×10^{-3}	1.81×10^{-3}	0.57
Annual	1.82×10^{-3}	2.00×10^{-3}	1.64×10^{-3}	0.57
		Maximum		
Winter	2.4×10^{-4}	4.8×10^{-4}	0	0.07
Spring	-8.6×10^{-4}	-5.8×10^{-4}	-11.4×10^{-4}	−0.22
Summer	-6.5×10^{-4}	-2.5×10^{-4}	-10.5×10^{-4}	−0.11
Fall	-2.7×10^{-4}	-0.5×10^{-4}	-4.9×10^{-4}	−0.18
Annual	-3.9×10^{-4}	-1.7×10^{-4}	-5.6×10^{-4}	−0.12
		Minimum		
Winter	3.18×10^{-3}	3.50×10^{-3}	2.86×10^{-3}	0.57
Spring	3.12×10^{-3}	3.41×10^{-3}	2.83×10^{-3}	0.60
Summer	4.00×10^{-3}	4.44×10^{-3}	3.56×10^{-3}	0.54
Fall	4.10×10^{-3}	4.48×10^{-3}	3.72×10^{-3}	0.60
Annual	3.61×10^{-3}	3.93×10^{-3}	3.37×10^{-3}	0.62
		Range		
Winter	-2.97×10^{-3}	-2.59×10^{-3}	-3.35×10^{-3}	−0.48
Spring	-3.98×10^{-3}	-3.55×10^{-3}	-4.41×10^{-3}	−0.54
Summer	-4.79×10^{-3}	-4.29×10^{-3}	-5.29×10^{-3}	−0.56
Fall	-4.36×10^{-3}	-3.88×10^{-3}	-4.84×10^{-3}	−0.54
Annual	-3.99×10^{-3}	-3.57×10^{-3}	-4.41×10^{-3}	−0.56

The first problem was related to the appropriate transformation of the population variable as represented by the urban metropolitan population. The regression equation used was of the form:

$$\bar{T}^*_{u-r} = f(\text{POPULATION}_u), \tag{1}$$

where population refers to the urban population and $\bar{T}^*_{u-r}$ is the mean urban minus rural temperature difference after the differences between the stations in elevation and latitude are considered. Several transformations of population were chosen; the square root and the logarithm of population were tested first. Mitchell (1953) used the square root of population and Oke (1973, 1976) the logarithm of population in his work dealing with instantaneous $\bar{T}_{u-r}$. Here, the square root of population proved more satisfactory, as statistically significant higher values of R in each of the separate regression equations were obtained using the 1950 and 1980 populations. As Mitchell (1953) argued, the

[2] This means that for a few urban pairs, the 1950 population was less than 2000.

square root function may have its basis in the tendency of the area of a city to: 1) be proportional to its population; 2) change in area proportional to its population change; and 3) be roughly circular in shape. The reason why the square root function proved more satisfactory than the logarithm in the regression equations is of interest. Oke (1976) has convincingly demonstrated the utility of using the logarithm of population, and more recently (Oke 1982), he has attached physical significance to the use of the logarithm of population in terms of city canyon sky-view factors. It must be remembered, however, that his work addresses the maximum observed urban-to-rural temperature difference, which is usually measured in parts of the city that have the greatest development, i.e., the lowest sky view factors. In this analysis, however, it is unlikely that an urban station will be located in that part of the city with the greatest development or lowest sky view factor. Instead, observing stations may be scattered across the city, and the probability that a station is located in an area with relatively low canyon sky view factor is likely to be disproportionately higher for cities with large populations and many developed areas than for cities with small populations and only one central developed area. In this regard, the square root function increases with population at a faster rate than the logarithm. This, then, could help explain the better relation of the square root function with heat islands studied herein. Furthermore, it would suggest that there is no reason why the square root of population should be the best dependent variable, and that other powers should be considered if the residuals of the regression equation are not well behaved. For both the 1950 and 1980 data, the resulting form of (1) then became

$$\bar{T}^*_{u-r} = a\,(\mathrm{POPULATION}_u)^b, \qquad (2)$$

where a is a coefficient determined by least squares methods and b equals 0.5.

The second problem concerned the variance of the residuals of Eq. (2) which were not random for predicted values of $\bar{T}^*_{u-r}$ and were heteroscedastic. This undesirable characteristic was addressed by a weighted least squares regression analysis as suggested by Draper and Smith (1966). For each of the three categories of population defined in Table 1, the relative importance of the observations within each category was made equal. Since there are three categories, the weight of each observation was defined by

$$0.333(N/N_c), \qquad (3)$$

where N is the total number of observations and N_c is the number of observations in category c.

In order to test whether the form of (2) was an appropriate *linear* equation, a was calculated for each of the three categories of urban population in Table 1 using unweighted least squares procedures. For these calculations, both the 1950 and the 1980 subsets were merged together, since the coefficients were quite consistent for these two subsets. After calculating a, it became apparent that, as the population increased, the value of a tended to become slightly smaller in absolute value, especially for the average temperature. Furthermore, several of the coefficients from each category were found to be outside of others' 95% confidence interval. Other values of b were tried in order to resolve this problem, i.e., b = 0.35, 0.40, and 0.45. A satisfactory result was achieved using

$$\bar{T}^*_{u-r} = a\,(\mathrm{POPULATION}_u)^{0.45}. \qquad (4)$$

Tables 2–5 contain the results of these tests with b = 0.45. Inspection of the coefficients indicates that all the coefficients, regardless of population category, are contained within the limits of others' 95% confidence interval. Table 6 was prepared in order to show any remaining systematic tendency of decreasing values of a with increasing population. None exists for the average or the maxima, but it could be argued that the exponent b = 0.45 has undercorrected for the minimum and temperature range. It is possible that a more appropriate exponent in (4) would be slightly below 0.45 for the range and the minimum, but the end points of the 95% confidence intervals for each of the population categories were quite close; therefore, no such change was made. Slightly different values of b for the temperature range and the minimum temperature

TABLE 3. As in Table 2, except for population of less than 10 000.

Season	Coefficient a	Coefficient 95% confidence interval: Upper	Lower	R
		Average		
Winter	1.46×10^{-3}	3.12×10^{-3}	-0.20×10^{-3}	0.08
Spring	2.03×10^{-3}	3.35×10^{-3}	0.71×10^{-3}	0.13
Summer	3.19×10^{-3}	5.06×10^{-3}	1.31×10^{-3}	0.15
Fall	1.73×10^{-3}	3.25×10^{-3}	0.21×10^{-3}	0.10
Annual	2.07×10^{-3}	3.43×10^{-3}	0.71×10^{-3}	0.13
		Maximum		
Winter	-1.41×10^{-3}	0.19×10^{-3}	-3.01×10^{-3}	−0.08
Spring	-0.74×10^{-3}	0.89×10^{-3}	-2.39×10^{-3}	−0.04
Summer	0.33×10^{-3}	2.39×10^{-3}	-1.72×10^{-3}	0.01
Fall	-1.17×10^{-3}	0.35×10^{-3}	-2.69×10^{-3}	−0.07
Annual	-0.77×10^{-3}	0.63×10^{-3}	-2.17×10^{-3}	−0.05
		Minimum		
Winter	4.11×10^{-3}	6.71×10^{-3}	1.51×10^{-3}	0.14
Spring	4.90×10^{-3}	7.10×10^{-3}	2.70×10^{-3}	0.20
Summer	6.75×10^{-3}	9.35×10^{-3}	4.15×10^{-3}	0.23
Fall	4.80×10^{-3}	7.60×10^{-3}	2.20×10^{-3}	0.15
Annual	5.12×10^{-3}	7.52×10^{-3}	2.72×10^{-3}	0.20
		Range		
Winter	-5.56×10^{-3}	-2.76×10^{-3}	-8.36×10^{-3}	−0.18
Spring	-5.65×10^{-3}	-2.65×10^{-3}	-8.65×10^{-3}	−0.17
Summer	-6.74×10^{-3}	-3.34×10^{-3}	-10.14×10^{-3}	−0.18
Fall	-5.97×10^{-3}	-2.57×10^{-3}	-9.37×10^{-3}	−0.16
Annual	-5.89×10^{-3}	-2.89×10^{-3}	-8.89×10^{-3}	−0.18

TABLE 4. As in Table 2, except for population of between 10 000 and 100 000.

Season	Coefficient a	Coefficient 95% confidence interval Upper	Lower	R
		Average		
Winter	1.48×10^{-3}	2.66×10^{-3}	0.30×10^{-3}	0.15
Spring	1.26×10^{-3}	2.20×10^{-3}	0.32×10^{-3}	0.16
Summer	2.13×10^{-3}	3.55×10^{-3}	0.71×10^{-3}	0.18
Fall	1.27×10^{-3}	2.27×10^{-3}	0.27×10^{-3}	0.15
Annual	1.51×10^{-3}	2.51×10^{-3}	0.61×10^{-3}	0.20
		Maximum		
Winter	-0.80×10^{-3}	0.60×10^{-3}	-2.00×10^{-3}	−0.08
Spring	-1.84×10^{-3}	0.32×10^{-3}	-3.36×10^{-3}	−0.14
Summer	-1.89×10^{-3}	0.31×10^{-3}	-4.09×10^{-3}	−0.10
Fall	-1.13×10^{-3}	0.13×10^{-3}	-2.39×10^{-3}	−0.11
Annual	-1.14×10^{-3}	0.16×10^{-3}	-2.44×10^{-3}	−0.13
		Minimum		
Winter	3.45×10^{-3}	5.15×10^{-3}	1.75×10^{-3}	0.24
Spring	3.26×10^{-3}	4.56×10^{-3}	1.96×10^{-3}	0.29
Summer	4.81×10^{-3}	6.37×10^{-3}	3.25×10^{-3}	0.35
Fall	3.09×10^{-3}	4.85×10^{-3}	1.33×10^{-3}	0.20
Annual	3.65×10^{-3}	5.07×10^{-3}	2.23×10^{-3}	0.30
		Range		
Winter	-4.33×10^{-3}	-2.61×10^{-3}	-6.05×10^{-3}	−0.29
Spring	-5.11×10^{-3}	-2.91×10^{-3}	-7.31×10^{-3}	−0.27
Summer	-6.94×10^{-3}	-4.34×10^{-3}	-9.54×10^{-3}	−0.30
Fall	-4.22×10^{-3}	-1.82×10^{-3}	-6.62×10^{-3}	−0.21
Annual	-5.08×10^{-3}	-3.08×10^{-3}	7.08×10^{-3}	−0.28

compared to the other elements could not be adequately justified on any physical basis. Furthermore, such a change would alter the results to only a very small degree.

The information provided in Tables 2–5 is of particular importance because it indicates that (4) can be applied to very low as well as to very high urban populations, a result that is not so obvious from an inspection of the scatter in Fig. 11, especially for low populations. In particular, the large number of urban-to-rural station pairs in the low population categories helps to offset the large variance in residuals and reduces what otherwise would be very large standard errors of estimate for a.

2) CHARACTERISTICS OF THE REGRESSION EQUATIONS

Some of the important characteristics of the regression equations are immediately apparent (Table 2). First, for most seasons the coefficients associated with the minima have a different sign than those associated with the maxima. Furthermore, the magnitude of the coefficients for the minima are much larger (by approximately an order of magnitude) than those for the maxima. These differing signs and the relatively large positive coefficients of the minima produce negative coefficients for the temperature range for all seasons. The magnitude of the coefficients for the average temperature is about one-half that for the minimum temperature, but the signs are unchanged.

The seasonal variation of the magnitude of the coefficients is significant. The coefficients are larger for the range, minima, and average during the summer and fall months than during the winter and spring months. The changes in the minima are mainly responsible for the larger coefficients of the average and range during summer and fall.

5. Interpretation of the results

A comparison of the results of the TCM and the TAM was completed by applying the equations derived from the TAM, in the same format as that used for the TCM, to predict the magnitude of the urbanization effect. Equation (4), the coefficients in Table 2, and twentieth-century population statistics were used to calculate the change in the $\bar{T}_{u-r}$ over the two sets of time periods, 1941–62 versus 1963–84 and 1901–42 versus 1943–84, using the station pairs depicted in Figs. 1–4 for each set. Population estimates were obtained

TABLE 5. As in Table 2, except for population greater than 100 000.

Season	Coefficient a	Coefficient 95% confidence interval Upper	Lower	R
		Average		
Winter	1.74×10^{-3}	2.12×10^{-3}	1.36×10^{-3}	0.81
Spring	1.46×10^{-3}	1.86×10^{-3}	1.06×10^{-3}	0.75
Summer	1.94×10^{-3}	2.84×10^{-3}	1.04×10^{-3}	0.53
Fall	1.86×10^{-3}	2.32×10^{-3}	1.40×10^{-3}	0.78
Annual	1.74×10^{-3}	2.18×10^{-3}	1.30×10^{-3}	0.71
		Maximum		
Winter	0.28×10^{-3}	0.88×10^{-3}	-0.32×10^{-3}	0.14
Spring	-0.68×10^{-3}	0.04×10^{-3}	-1.40×10^{-3}	−0.28
Summer	-0.52×10^{-3}	0.48×10^{-3}	-1.52×10^{-3}	−0.16
Fall	-0.24×10^{-3}	0.24×10^{-3}	-0.72×10^{-3}	−0.16
Annual	-0.29×10^{-3}	0.25×10^{-3}	-0.83×10^{-3}	−0.16
		Minimum		
Winter	3.02×10^{-3}	3.76×10^{-3}	2.28×10^{-3}	0.78
Spring	3.02×10^{-3}	3.82×10^{-3}	2.22×10^{-3}	0.75
Summer	3.81×10^{-3}	5.17×10^{-3}	2.45×10^{-3}	0.65
Fall	3.79×10^{-3}	4.79×10^{-3}	2.79×10^{-3}	0.75
Annual	3.42×10^{-3}	4.32×10^{-3}	2.42×10^{-3}	0.76
		Range		
Winter	-2.77×10^{-3}	-1.67×10^{-3}	-3.87×10^{-3}	−0.61
Spring	-3.71×10^{-3}	-2.51×10^{-3}	-4.91×10^{-3}	−0.68
Summer	-4.41×10^{-3}	-3.71×10^{-3}	[illegible]	−0.74
Fall	-4.02×10^{-3}	-2.76×10^{-3}	-5.28×10^{-3}	−0.69
Annual	-3.71×10^{-3}	-2.55×10^{-3}	-4.87×10^{-3}	−0.70

TABLE 6. Relative position of the magnitude of the coefficient a of the regression equation: $\bar{T}^*_{u-r} = a\,(\text{population}_u)^{0.45}$ with respect to three categories of population: 1) 2000 to 10 000, 2) 10 000 to 100 000, and 3) >100 000. Here L, M, H, stand for the rank of the coefficients: low, medium, or high.

Season	1	2	3
	Average		
Winter	L	M	H
Spring	H	L	M
Summer	H	M	L
Fall	M	L	H
Annual	H	L	M
	Maximum		
Winter	L	M	H
Spring	M	H	L
Summer	L	H	M
Fall	H	M	L
Annual	M	H	L
	Minimum		
Winter	H	M	L
Spring	H	M	L
Summer	H	M	L
Fall	H	L	M
Annual	H	M	L
	Range		
Winter	H	M	L
Spring	H	M	L
Summer	M	H	L
Fall	H	M	L
Annual	H	M	L

from the U.S. Census Bureau statistics for the 1920, 1950, and 1980 census. From these census figures the population of each station during each year from 1900 to 1984 was estimated by using the time rate of change derived from the difference in population 1950–20 for the period prior to 1950, and 1980–50 for the period after 1950. Figures 9 and 10 give the results of such calculations for the TAM. Several general characteristics are apparent. First, for the TAM, the difference between the rate of urban warming (or cooling) is consistent regardless of the time period (either 1901–84 or 1941–84) chosen for comparison. For example, for the average temperature, the warming rate of the most populous category is approximately 0.09°C decade^{-1} for the period from 1941 to 1984 and 0.08°C decade^{-1} for the whole period from 1901 to 1984. This consistency in the rates of change holds for all elements and population categories. The same cannot be said for the TCM results. The warming due to urbanization for the maximum temperature is greater by nearly 0.04°C for two of the three population categories solely by changing time periods. The results from the TAM indicate that this cannot be attributed to a change in the network of station pairs nor to a large change in the rate of population increase. It is more likely that this variability is a result of the noise in the time series introduced by imperfect adjustments for discontinuities, undocumented inhomogeneities, changes in the frequency of air mass regimes or weather types, and other inherent random and nonrandom variations.

The primary difference between the two methods of deriving average temperature is manifested in the highest population category (>100 000). Compared to the TCM, the TAM predicts nearly twice the bias due to urban warming. Both methods, depict a positive urban bias, however, even at very low urban populations. This is consistent with the results of Landsberg (1975) for Columbia, Maryland. The difference in the highest category (>100 000) is due to the much larger warming rate of the TAM for the minimum temperature. In the same category, very large differences occur in the rate of temperature decrease of the diurnal temperature range. There is much closer agreement for the two lower population categories with respect to the sign and magnitude of the rates of change of $\bar{T}_{u-r}$ for most elements. In these two categories, only during the period from 1941 to 1948 is there a discrepancy between the two methods, namely, in the maximum temperature for the 10 000 to 100 000 category.

In terms of seasonal variations, both the TAM and the TCM indicate that the rate of warming for the minimum and the average temperature is larger during summer and fall than during winter and spring. This is consistent with the results of other work for both $\bar{T}_{u-r}$ and the instantaneous maximum T_{u-r} in much of the United States (Landsberg 1981). The season with the smallest urban effect is spring. Apparently, the relative infrequency of calm nights in midlatitudes during the spring as compared to the other seasons, and the smaller amounts of anthropogenic heat compared to winter, prevent the urbanization effect in spring from being as large as that during other seasons.

6. Implications of urbanization

The results of previous studies can be used to help resolve the differences between the TCM and the TAM. Prior analyses, which have addressed both the maximum and minimum temperature using variations of either the TCM (Cayan and Douglas 1984) or the TAM (Chandler 1965; Landsberg 1981; Hage 1972; Garstang et al. 1975), have all indicated that the minimum temperature is more strongly affected by urbanization than the maximum. Furthermore, the many studies of instantaneous T_{u-r} suggest that the urbanization effects in daytime (maximum temperature) are small; some studies have detected a "cool island" (Oke 1982). This is consistent with the results of the TAM. Because of this, and because

1) the TCM is somewhat inconsistent from time period to time period with respect to the warming rate of the maximum temperature;

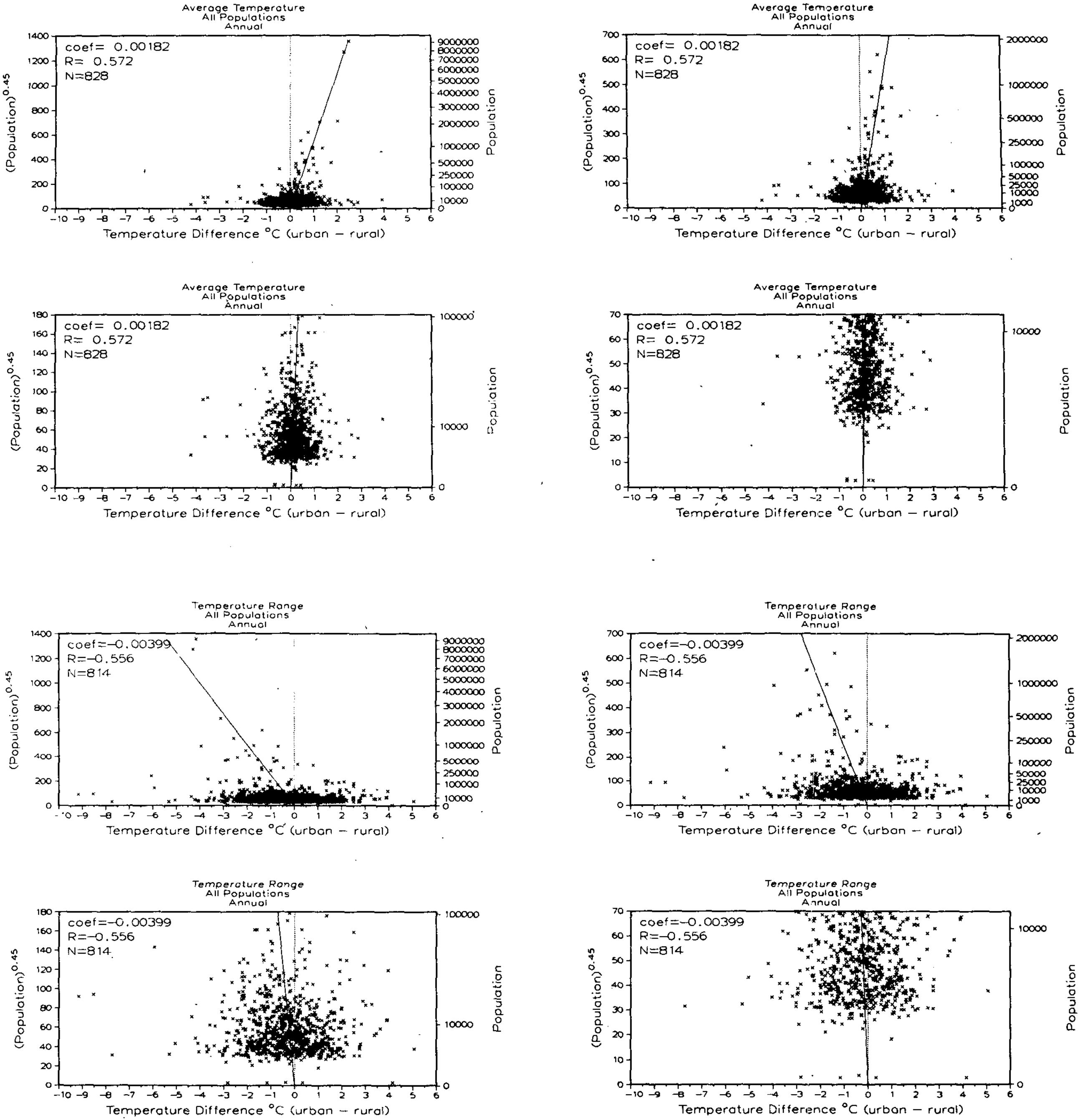

FIG. 11. Scatter plots of $\bar{T}^{*}_{u-r}$ (see section 3) related to population shown in four different ranges. The solid line depicts regression line of best fit, and a vertical dashed line represents no relation between population and $\bar{T}^{*}_{u-r}$. The regression line is based on all urban populations for 1950 and 1980 combined.

2) the TCM is vulnerable to imperfect adjustments for station changes;

3) the TCM can be adversely affected by undocumented changes in the local landscape in the immediate vicinity of the station any time over the entire twentieth century;

4) the TCM can be affected by changes in the frequency of air mass regimes over time; and

5) there is uncertainty about the most appropriate technique for calculating the time rate of change due to the fluctuations and vacillations in the series; the results from the TAM are preferred.

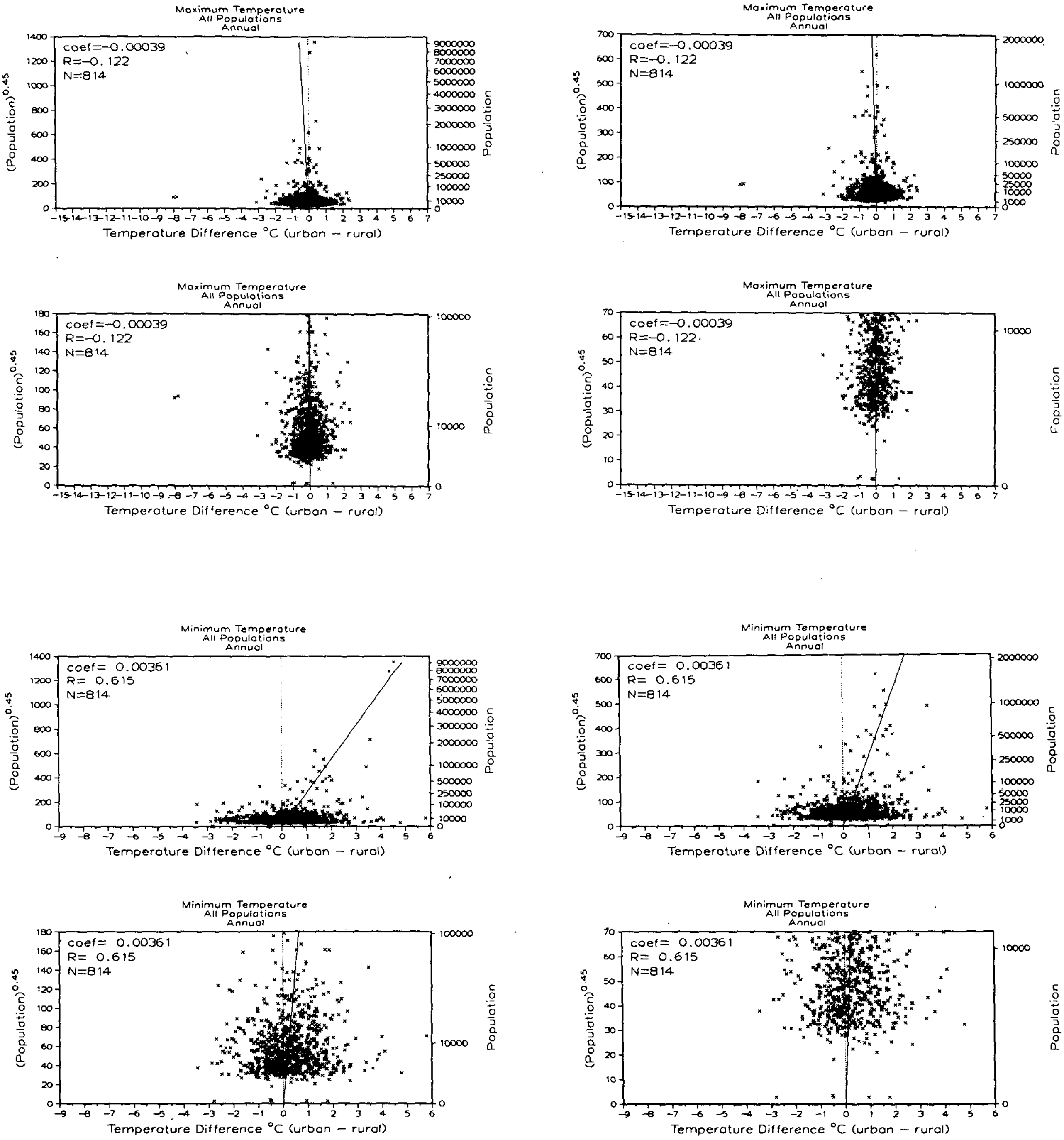

FIG. 11. *(Continued)*

The TAM was used to calculate the effect of urbanization on the climate record of the United States in the HCN. Stations from two subsets of the original 1219 stations were used. In both subsets, data were adjusted for potential discontinuities. In one subset (subset *a*), at least 75% of the adjusted data had to be present (no more than 25% of the data estimated), and the stations with the widest confidence intervals in the adjustments were not used (i.e., the widest 10% were omitted). In the other subset (subset *b*), all stations that could be adjusted were used, regardless of quality. Area-weighted temperatures were calculated by forming 23 regions in the United States (Fig. 12) and identifying a fixed network of stations for various time periods for

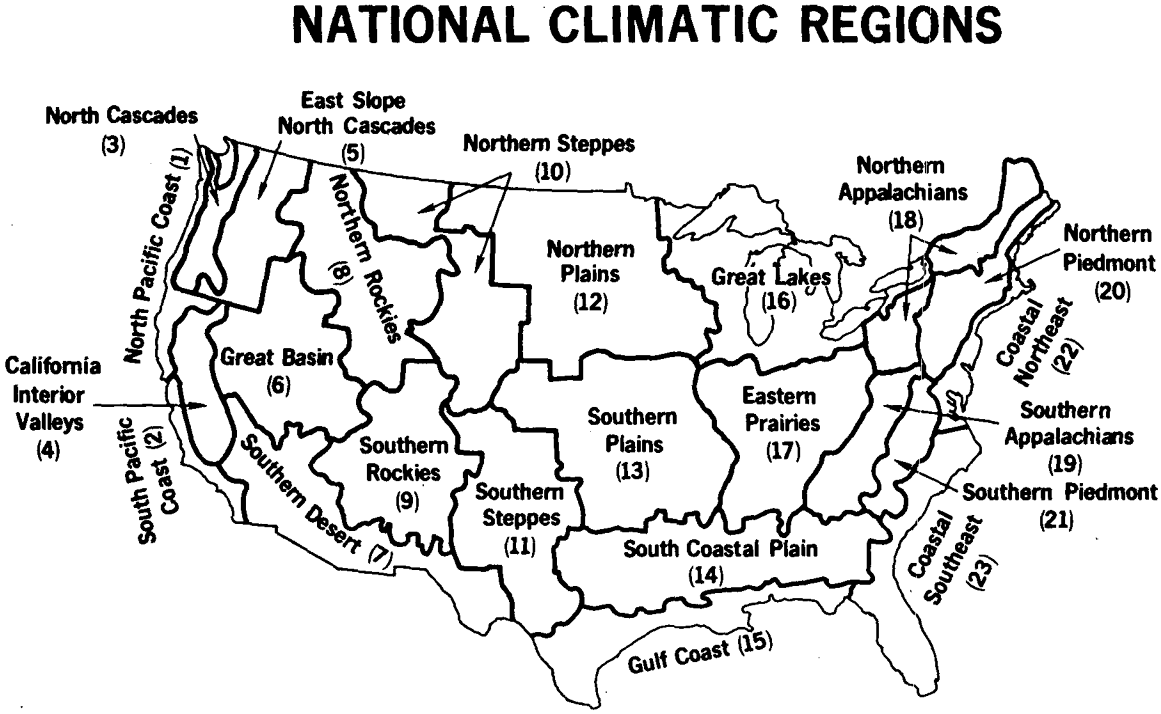

FIG. 12. Regions used to area weight temperatures in the United States area-average calculation.

the maximum, minimum, and range, as well as for the average temperature (Fig. 13). Seasonal means for each region were calculated by weighting each station within a region equally, and a United States area mean was calculated by weighting each area according to the area of the region relative to the area of the contiguous United States. Adjustments for urbanization effects were applied to the data using (4) for each network and for each season. The populations of each station were derived as described in section 5. Figure 14 depicts the urbanization corrected and noncorrected temperature time series for the "higher" quality network (subset *a*). The urbanization bias in these networks is small. It averages approximately +0.06°C for the average temperature, +0.13°C for the minimum temperature, −0.01°C for the maximum temperature, and a −0.14°C for the temperature range over the 1901–84 time period. When the same urbanization corrections are applied to the subset *b* time series, the heat island bias is nearly identical in both networks.

Much of what is presented in Fig. 14 with respect to the mean temperature has been noted by others, who have used mostly nonurban stations (Diaz and Quayle 1980; Karl et al. 1984), but the changes in the maximum, minimum, and temperature range in the United States over the entire twentieth century have never before been documented. One of the most interesting characteristics of the climate record is the decrease in the diurnal temperature range since the 1940s and the lack of any noticeable trend prior to this time. Karl et al. (1984, 1986b, 1987) discuss this recent decrease as it may relate to the greenhouse effect or aerosol loading.

Table 7 and Fig. 15 provide an estimate of the effect of urbanization for cities of various sizes compared to a rural station. It is clear from Figs. 14 and 15 that, due to the large number of rural stations in the HCN, the effect of population growth in the HCN networks is not large when temperature data are aggregated across much of the HCN database. The average difference in the mean temperature trend is only 0.06°C, quite small compared to the year-to-year variability of the annual mean and many of the multiyear climate fluctuations or changes in level (cf. Fig. 14). However, the use of another dataset with a greater urban representation, if uncorrected, could easily result in much larger biases.

7. Conclusion and recommendations

Urbanization has influenced the climate records of even small towns in the United States (Fig. 15). A method has been developed that, when used with large-scale area averages, removes most of the bias for the temperature maximum, minimum, average, and range, given the history of station population. The urbanization bias was found to be a predominately nighttime phenomenon especially pronounced in minimum temperature and the diurnal temperature range.

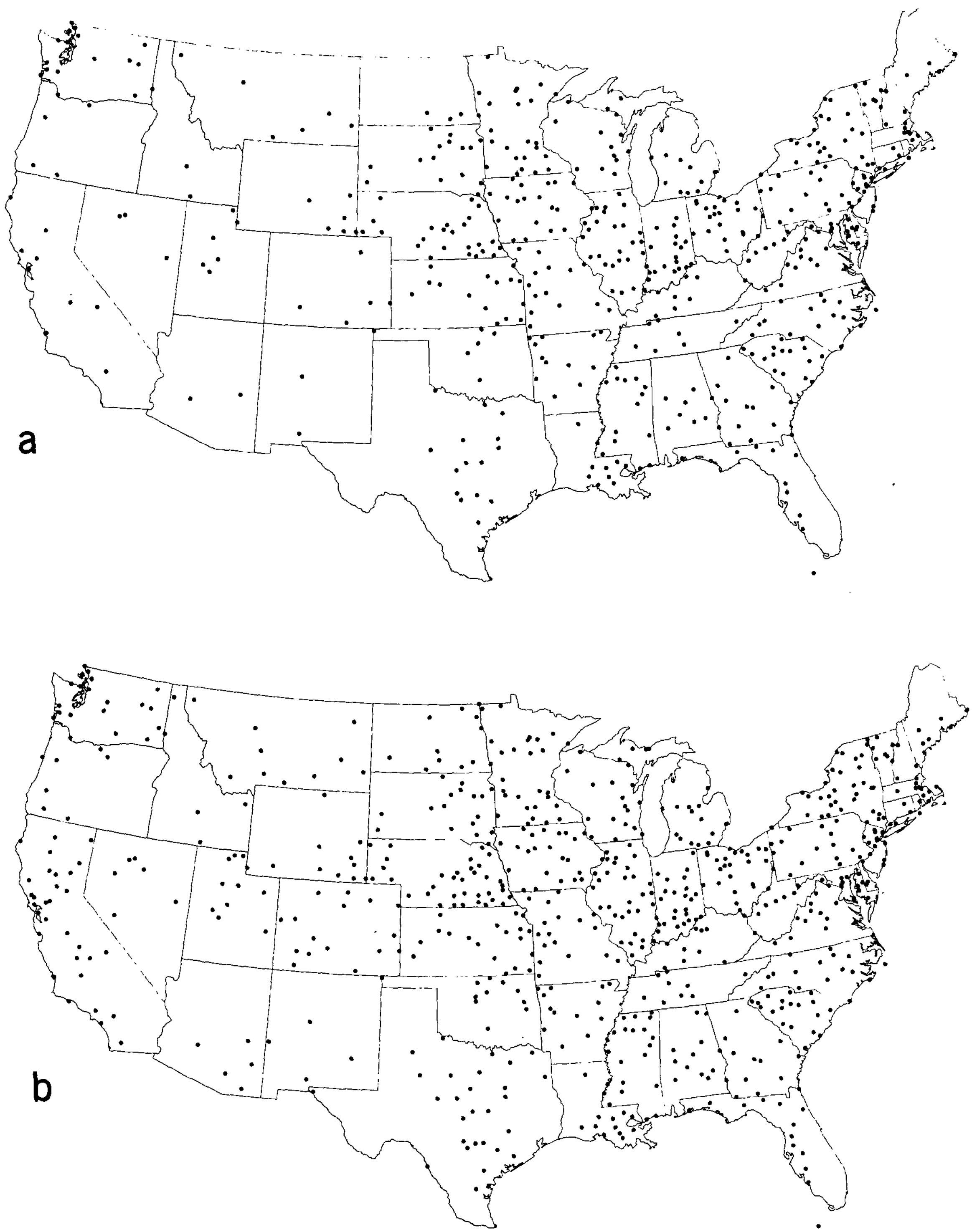

FIG. 13. Fixed network of stations (1901–84) where the confidence of the adjustments and the amount of estimated data (a) were considered and (b) where they were not considered.

The effect of urbanization on the Historical Climatology Network during the twentieth century (1901–84) is shown to be practically negligible for the maximum temperature but significant for the diurnal minima, means, and range. On an annual basis it amounts to 0.06°C, 0.13°C and −0.14°C for the average, min-

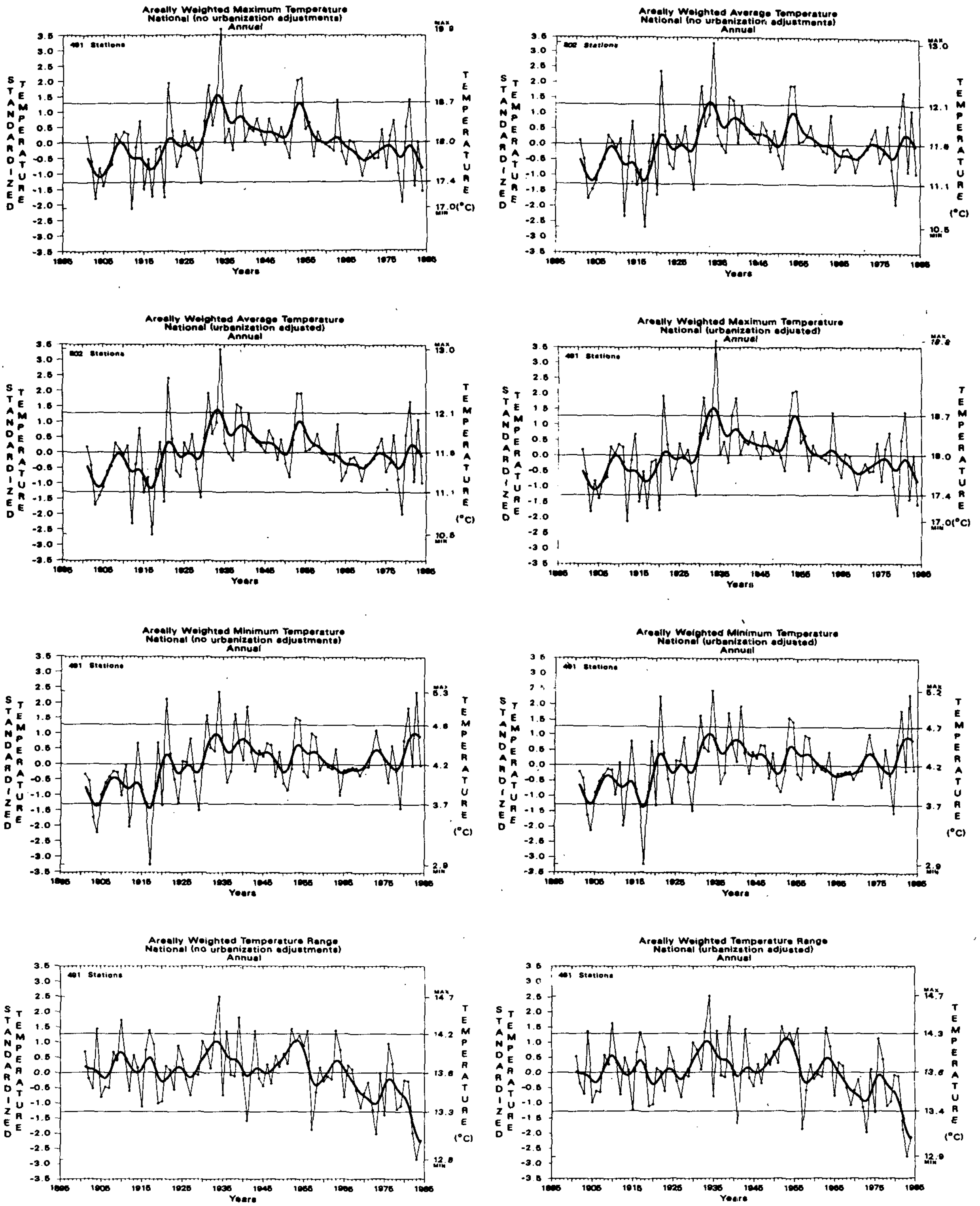

FIG. 14. Time series of urban adjusted (right panel) versus unadjusted (left panel) area-weighted mean temperature for the United States for the higher-quality stations in the Historical Climatology Network (Fig. 13a).

TABLE 7. The effect of urbanization on the annual temperature (°C) in the United States with respect to a station located in a rural environment (population < 2000). Each element is derived independently of other elements [(max + min)/2 = avg].

Population	Average	Maximum	Minimum	Diurnal temperature range
2000	0.06	−0.01	0.12	−0.13
5000	0.08	−0.02	0.16	−0.18
10 000	0.11	−0.02	0.22	−0.24
20 000	0.16	−0.03	0.32	−0.35
50 000	0.24	−0.05	0.48	−0.53
100 000	0.32	−0.07	0.63	−0.70
200 000	0.44	−0.09	0.87	−0.96
500 000	0.67	−0.14	1.33	−1.47
1 000 000	0.91	−0.20	1.81	−2.00
2 000 000	1.25	−0.27	2.48	−2.74
5 000 000	1.88	−0.40	3.73	−4.12
10 000 000	2.57	−0.55	5.10	−5.63

imum and diurnal temperature range, respectively. It must be noted, however, that the Historical Climatology Network is a relatively rural station network, in which over 85% of the stations are located in cities with a population of less than 25 000 in 1980. This implies that the growth rate of these stations has been quite small over the twentieth century. Results from subsets of the network and from other studies indicate that much larger biases are present in rapidly growing urban areas. Several additional points about the results deserve emphasis:

1) The results are based on a statistical aggregate over many station pairs.

2) There are strong fluctuations in long time series of urban-to-rural temperature differences, which may be related to climate variability or local changes around each site. This makes it difficult to assess the effect of urbanization using such time series, especially for stations with small urban warming biases.

3) No significant regional differences were detected, but some significant seasonal differences were found with respect to the magnitude of the urban bias. The bias was found to be greatest during summer and fall, and least during winter and spring.

4) The adjustment scheme used in the HCN had little effect on the results.

Some words of caution are pertinent. The rural stations used to estimate the impact of urbanization had an average population in 1980 of 700, so that, for a station located in an unpopulated area, the computed magnitude of the annual average heat island effect is underestimated. On the other hand, if the equations we developed are extrapolated to a population of zero, such underestimates are likely to be relatively small, no more than 0.03°C.

We do not recommend the use of the equations developed in this study to predict the impact of a heat island at any particular station as the explained variance of the equations is not high enough for such an appraisal to be valid. Instead, the equations should be used only if a number of stations are available to represent regional averages. Local climate effects, the specific location of the station in relation to the city center, and the special attributes of any given city, possibly unrelated to population, preclude the prediction of the urban effect at any single location by this methodology alone.

For similar reasons, the general application of such equations in a global context cannot be recommended

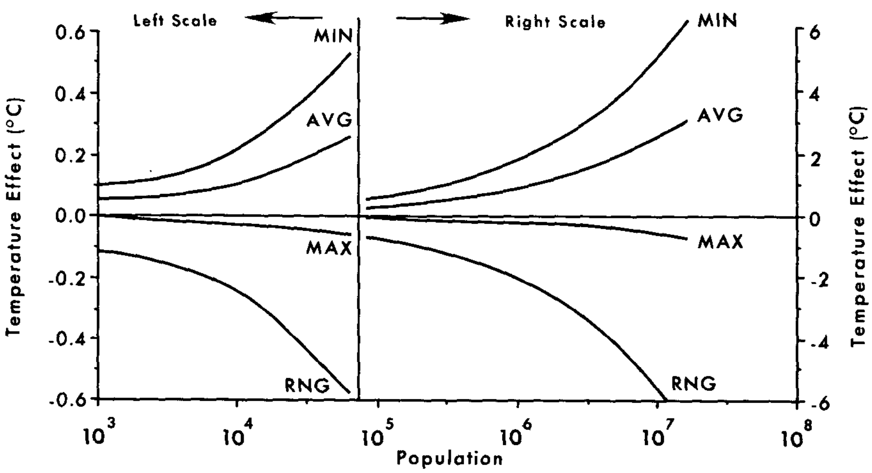

FIG. 15. Annual mean temperature bias due to urban effects as a function of population derived from stations in the U.S. Historical Climatology Network. The left scale is for population up to 90 000 and the right scale for populations of 90 000 or more. Here AVG implies average temperature, MAX, maximum temperature, MIN, minimum temperature, and RNG, temperature range.

TABLE 8. Predicted annual T_{u-r} (°C) using Oke's (1976) equations.

Population	Europe	North America	Ratio
2000	2.6*	3.3	0.79
5000	3.4*	4.5	0.76
10 000	4.0*	5.5	0.73
20 000	4.6*	6.4	0.72
50 000	5.4	7.6	0.71
100 000	6.0	8.5	0.71
200 000	6.6	9.4	0.70
500 000	7.4	10.6	0.70
1 000 000	8.0	11.6	0.69
2 000 000	8.6	12.5	0.69
5 000 000	9.4*	13.7	0.69
10 000 000	10.0*	14.6	0.68

* Predicted values are outside the bounds of the developed regression equations.

either. The varying ratios of anthropogenic heat generated by the city to amount of insolation received, quantities of evapotranspiration, ventilation, etc., are likely to be appreciably different in tropical and polar climates compared to the midlatitudes. Indeed, prolonged intense inversions in polar climates have been shown to produce very large heat islands (Weller 1982). On the other hand, the climate of the United States may be similar enough to other midlatitude locations in Europe and Asia that the equations developed might be effective. Oke (1976, 1982) provides evidence to suggest, however, that the estimate of the greatest instantaneous T_{u-r} for any given city is characteristically different for European and North American cities because of characteristic differences in canyon sky-view factors for equal populations. Table 8 contains the estimated effect of population on the largest instantaneous T_{u-r} in Europe and North America. Since these values were associated with clear, calm nocturnal temperatures, one can assume that a large portion of the differences can be attributed to the characteristics of the cities themselves. If so, then one could conceivably use the ratio of the maximum instantaneous T_{u-r} of European versus North American cities as an adjustment to the predicted effect of urbanization as derived from the equations developed in this study, but this should only serve as an interim solution to the problem until data can be examined in other parts of the world. At the present time, however, the station density of computer compatible global land surface datasets needs improvement.

Acknowledgments. This work was supported by the Carbon Dioxide Research Division of the U.S. Department of Energy Contract No. AC02-DE-FG02-85ER60372. We thank Dr. Tim Oke for his valuable criticism and Dr. Dan Cayan and Robert Quayle for their constructive comments. Although we have taken their advice on many points, there may still remain some areas of disagreement.

REFERENCES

Cayan, D. R., and A. V. Douglas, 1984: Urban influences on surface temperatures in the southwestern United States during recent decades. *J. Climate Appl. Meteor.*, **23**, 1520–1530.
Chandler, T. J., 1965: *The Climate of London*, Hutchinson, 150 pp.
Chow, S. D., 1986: Some aspects of the urban climate of Shanghai. *Urban Climatology and Its Applications with Special Regard to Tropical Areas.* World Meteorological Organization, No. 652, 87–109.
Clarke, J. F., 1969: Nocturnal urban boundary layer over Cincinnati, Ohio. *Mon. Wea. Rev.*, **97**, 582–589.
Dettwiller, J., 1970: Deep soil temperature trends and urban effects at Paris. *J. Appl. Meteor.*, **9**, 178–180.
Diaz, H. F., 1986: An analysis of twentieth century climate fluctuations in northern North America. *J. Climate Appl. Meteor.*, **25**, 1625–1657.
——, and R. G. Quayle, 1980: The climate of the United States since 1895: spatial and temporal change. *Mon. Wea. Rev.*, **108**, 249–266.
Draper, N. R., and H. Smith, 1966: *Applied Regression Analysis*, John Wiley and Sons, 407 pp.
Duchon, C. E., 1986: Temperature trends at San Juan, Puerto Rico. *Bull. Amer. Meteor. Soc.*, **67**, 1370–1377.
Ellsaesser, H. W., M. C. MacCracken, J. J. Walton and S. L. Grotch, 1986: Global climatic trends as revealed by the recorded data. *Rev. Geophys.*, **24**, 745–792.
Emerson, J. D., and D. C. Hoaglin, 1983: Resistant lines for *y* versus *x*. *Understanding Robust and Exploratory Data Analysis.* D. C. Hoaglin, F. Mosteller and J. W. Tukey, Eds. John Wiley and Sons, 129–165.
Fukui, E., 1970: The recent rise of temperature in Japan. *Japan. Progr. Climatol.*, Tokyo University of Education, 46–65.
Garstang, M., P. D. Tyson and G. D. Emmitt, 1975: The structure of heat islands. *Rev. Geophys. Space Phys.*, **13**, 139–165.
Goodridge, J. D., 1985: Urban and rural differences in long range temperature trends in California. *Extended Abstracts, Third Conf. on Climate Variations and Symp. on Contemporary Climate: 1850–2100*, Los Angeles, Amer. Meteor. Soc., 158–159.
Hage, K. D., 1972: Nocturnal temperatures in Edmonton, Alberta. *J. Clim. Appl. Meteor.*, **11**, 123–129.
Jones, P. D., S. C. B. Raper, R. S. Bradley, H. F. Diaz, P. M. Kelly and T. M. L. Wigley, 1986: Northern Hemisphere surface air temperature variations, 1851–1984. *J. Climate Appl. Meteor.*, **25**, 161–179.
Karl, T. R., 1988: Multi-year fluctuations of temperature and precipitation: The gray area of climate change. In press, *Climatic Change.*
——, and C. N. Williams, Jr., 1987: An approach to adjusting climatological time series for discontinuous inhomogeneities. *J. Climate Appl. Meteor.*, **26**, 1744–1763.
——, and P. J. Young, 1987: The 1986 Southeast drought in historical perspective. *Bull. Amer. Meteor. Soc.*, **68**, 773–778.
——, G. Kukla and J. Gavin, 1984: Decreasing diurnal temperature range in the United States and Canada, 1941–80. *J. Climate Appl. Meteor.*, **23**, 1489–1504.
——, C. N. Williams, Jr., P. J. Young and W. M. Wendland, 1986a: 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.
——, G. Kukla and J. Gavin, 1986b: Relationship between decreased temperature range and precipitation trends in the United States and Canada, 1941–80. *J. Climate Appl. Meteor.*, **26**, 1878–1886.
——, —— and ——, 1987: Recent temperature changes during overcast and clear skies in the United States. *J. Climate Appl. Meteor.*, **27**, 698–711.
Kukla, G., J. Gavin and T. R. Karl, 1986: Urban warming. *J. Climate Appl. Meteor.*, **25**, 1265–1270.
Landsberg, H. E., 1975: Atmospheric changes in a growing com-

munity. Inst. Fluid Dynamics Appl. Math. Tech. Note No. BN 823, Univ. of Maryland, 54 pp.
——, 1981: *The Urban Climate.* Academic Press, 285 pp.
Lowry, W. P., 1977: Empirical estimation of urban effects on climate: A problem analysis. *J. Appl. Meteor.,* **16,** 124–135.
Mitchell, J. M., Jr., 1953: On the causes of instrumentally observed secular temperature trends. *J. Meteor.,* **10,** 244–261.
——, 1961: The temperature of cities. *Weatherwise,* **14,** 224–227.
Oke, T. R., 1973: City size and the urban heat island. *Atmos. Environ.,* **7,** 769–779.
——, 1976: The distinction between canopy and boundary-layer urban heat islands. *Atmosphere,* **14,** 268–277.
——, 1979: Review of Urban Climatology. WMO Tech. Note No. 169, 100 pp.
——, 1981: Canyon geometry and the nocturnal urban heat island: Comparison of scale model and field observations. *J. Climatol.,* **1,** 237–253.
——, 1982: The energetic basis of the urban heat island. *Quart. J. Roy. Meteor. Soc.,* **108,** 1–24.
Quinlan, F. T., T. R. Karl and C. N. Williams, Jr., 1987: United States Historical Climatology Network (HCN) serial temperature and precipitation data. NDP-019, Carbon Dioxide Information Analysis, Oak Ridge National Laboratory, Oak Ridge, Tennessee. 33 pp. plus appendices.
Ropelewski, C. F., and M. S. Halpert, 1986: North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Mon. Wea. Rev.,* **114,** 2352–2362.
Weller, G., 1982: The urban climate in Alaska. *Annual Report 1981–82,* Geophysical Institute, University of Alaska, 1–22.
Wigley, T. M. L., J. K. Angell and P. D. Jones, 1985: Analysis of the temperature record. *Detecting the Climate Effects of Increasing Carbon Dioxide.* M. C. MacCracken and F. M. Luther, Eds. DOE/ER 0235, NTIS-PR-360, 55–90.
World Meteorological Organization, 1986: *Proceedings of the Technical Conf. on Urban Climatology and Its Applications with Special Regard to Tropical Areas.* Mexico City, 1984, WMO No. 652, 534 pp.
Yarnal, B., and H. F. Diaz, 1986: Relationships between extremes of the southern oscillation and the winter climate of the Anglo-American Pacific Coast. *J. Climatol.* **6,** 197–219.