Theor. Appl. Climatol. 79, 165–174 (2004)
DOI 10.1007/s00704-004-0065-4

Theoretical and Applied Climatology
Printed in Austria

[1] National Meteorological Center, CMA, Beijing, China
[2] Physical Institute, Peking University, Beijing, China
[3] Jiangsu Province Key Laboratory of Meteorological Disaster and Environmental Variation, Nanjing, China

Urban heat island effect on annual mean temperature during the last 50 years in China

Q. Li[1,3], **H. Zhang**[1], **X. Liu**[1], and **J. Huang**[2]

With 4 Figures

Received January 16, 2004; revised March 10, 2004; accepted April 4, 2004
Published online October 14, 2004

Summary

Based on China's fifth population survey (2000) data and homogenized annual mean surface air temperature data, the urban heat island (UHI) effect on the warming during the last 50 years in China was analyzed in this study. In most cities with population over 10^4, where there are national reference stations and principal stations, most of the temperature series are inevitably affected by the UHI effect. To detect the UHI effect, the annual mean surface air temperature (SAT) time series were firstly classified into 5 subregions by using Rotated Principal Components Analysis (RPCA) according to its high and low frequency climatic change features. Then the average UHI effect on each subregion's regional annual mean STA was studied. Results indicate that the UHI effect on the annual mean temperatures includes three aspects: increase of the average values, decrease of variances and change of the climatic trends. The effect on the climatic trends is different from region to region. In the Yangtze River Valley and South China, the UHI effect enhances the warming trends by about 0.011 °C/decade. In the other areas, such as Northeast, North-China, and Northwest, UHI has little impact on the warming trends of the regional annual temperature; while in the Southwest of China, introducing UHI stations slows down the warming trend by −0.006 °C/decade. But no matter what subregion it is, the total warming/cooling of these effects is much smaller than the background change in regional temperature. The average UHI effect for the entire country, during the last 50 years is less than 0.06 °C, which agrees well with the IPCC (2001). This suggests that we cannot conclude that urbanization during the last 50 years has had much obvious effect on the observed warming in China.

1. Introduction

The Urban Heat Island (UHI) effect has received much attention in climatic research (Mitchell, Jr., 1961). Oke (1973) concluded that even a city of 1000 people could have an UHI effect, and the magnitudes of the UHI effects were linearly correlated with the logarithms of the population. IPCC (2001) pointed out that there are two primary reasons why urban heat islands have been suspected as being partially responsible for the observed increases in land air temperatures over the last few decades. The first one is related to the observed decrease in the diurnal temperature range and the second is related to a lower rate of warming observed over the past twenty years in the lower troposphere compared with the surface; but at the same time it is minor compared with the background change of temperature, and the urban heat island effects are no more than about 0.05 °C up to 1990 in the global temperature records. Two different approaches were used by Karl et al. (1988) in an attempt to determine the effect of urbanization in the U.S. climate records:

one is the rate of time change method and the other one is an approach that analyzes mean urban–rural differences for two 9-year periods of urban–rural pairs. He found that the magnitudes of the UHI effect reach 0.11 °C in a city of 10^4 people, 0.32 °C in a city of 10^5 people and above 0.91 °C in a city of 10^6 people (Karl et al., 1988). Kanay and Cai (2002) estimated the impact of urbanization and other land uses on climate change by comparing trends observed at surface stations with surface temperature derived from the NCEP-NCAR 50 Reanalysis. Their estimate is 0.27 °C/century, which is at least twice as high as previous estimates based on urbanization alone. Peterson (2003) firstly applied a variety of adjustments to the temperature data to remove the biases caused by changes in elevation, latitude, time of observation, and instrumentation. After the adjustments, he found no statistically significant impact between urban and rural stations' surface temperature. For different countries and regions, the UHI effect will be different due to the situation of meteorological observations in different countries. Wang et al. (1990) chose 42 pairs of China urban (average population is $1.7 \cdot 10^6$) and rural (average population $1.5 \cdot 10^5$) stations to study the Urban Heat Island effects. However, some questions remained to be answered: Can these 84 stations represent the situation of the whole country? How much do the UHI effects affect the regional warming? And is a "rural" station with an average population of $1.5 \cdot 10^5$ not affected by UHI effect? Obviously, it is necessary to do further research to provide some objective assessment to these questions.

2. Data and homogeneity adjustment

As we noted before, the homogeneity and reliability of the climatic data is very important for UHI effect analysis. However, the *in situ* climatic data series that people use may be inhomogeneous due to relocations of stations, changes in instrumentation and daily mean temperature computing methods etc., so it is easy to obtain errors. In UHI effect studies, this is a major problem that causes biases (Portman, 1993). We note that many stations, especially some big city stations, have been relocated since 1951. For example, some urban stations like Beijing and Shanghai have been relocated several times between the suburbs and urban sites. So if inhomogeneity problems are not handled properly, it is impossible to obtain satisfactory research results. There have been many studies on the techniques of homogenization of climatic series (Peterson et al., 1998). In China, the homogeneity study of the climatic data is very limited (Zhai, 1996; Liu and Li, 2003). Recently, based on the previous studies, we have tested and adjusted the inhomogeneity of China's monthly temperature data and created a set of homogeneous monthly temperature data for the recent 51 years (1951–2001). In this dataset, the main discontinuities resulted from station relocations. Other factors were also detected, and the corresponding inhomogeneous temperature series have been adjusted (Li et al., 2004).

In the following we take Beijing as an example to illustrate our method of adjusting. Since 1951, Beijing station (54511) has been relocated 6 times, in 1953, 1965, 1969, 1970, 1981 and 1997. By applying Easterling-Peterson's techniques we found 5 discontinuities, which are in 1955, 1964, 1970, 1980, and 1996. Taking the latest relocation as an example, in 1996 the station moved from downtown to the current suburb location, where the annual temperature is obviously lower than the observed one before 1996. The adjusted temperature series is based on the present location, so the data in the period that the station was located in the urban site should be adjusted (Fig. 1). It can be clearly seen that the adjusted temperature series is

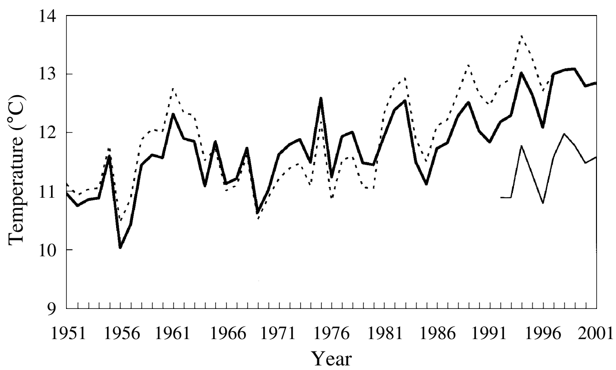

Fig. 1. Beijing's annual mean temperature time series before (dashed curve) and after (solid curve) adjustment. The short curve is annual temperature for Miyun (short curve), a rural station near Beijing

obviously different from the observational series in certain periods but it is very reasonable compared with the Miyun station, which lies in the rural sites and has not been relocated for the last 10 years. So we can conclude that after removing the relocation effects, we can get more reliable results.

According to above analysis, the behavior of the UHI effect on China's surface air temperature includes two aspects. Firstly, it is obviously warmer in urban sites than in the rural/suburb sites, so it is not appropriate to use temperature series from urban stations as the regional temperature change. On the other hand, some of the suburb stations have been changed into urban stations because of urbanization, thus leading to a certain amount of bias when calculating climate change on decadal scales. The ordinary method is to compare the time series of an urban station with that of a nearby rural station and then attribute the differences between them as the UHI effect. But as we know, there is much difficulty in finding a pure "rural" station near each "urban" one; so it always leads to randomness in the research results. In this paper, we are trying to find the average effects on whole regions where the annual temperature anomalies are similar for all stations. That is, how much the warming trend changed before and after the UHI stations are introduced will be regarded as the UHI effect on the regional temperature trend, which is easier to understand. So we first divide China's annual mean temperature field into several subregions and then analyze the impact of the UHI in each subregion.

For the convenience of calculation and analysis, we select about 390 annual mean temperature series that are comparatively long and spatially homogeneous from the reference and principal stations. The length of time series is 48 years, from 1954 to 2001, and some individual missing values were interpolated by comparing with the surrounding stations.

Traditionally, the urbanization can be described as function of population density (Karl et al., 1981; Peterson, 2003) and urban area, among which the population factor has certain advantage in obtaining data. Recently many researchers have introduced the lights at night observed by satellites to define urbanization (Gallo et al., 1996, 1999; Hansen et al., 2001), which obviously is an objective way for some countries and areas and it can determine the urbanization of a certain city more specifically. In China, due to the unbalanced economic development, living style, the awareness of saving energy and other factors, this method is not very suitable. Therefore, urban population is still used as the main index to determine the scale of the city where meteorological stations are located in our study. We use the city population data as the index of the magnitudes of the urbanization. The population data are compiled by National Bureau of Statistics of China during China's fifth population survey (2000), from which we selected the number of permanent inhabitants in a city as the index of the urbanization.

There have been many studies on classifying climate fields by using RPCA. The experience in RPCA applications (Horel, 1981; Li and Tu, 2002) shows that the selection of the number of principal components (PCs) for rotation is the key for climate classification. If too many PCs are rotated, there will be too few extremum centers and the central values will be too large, thus the local features are overemphasized (If the number of truncation points equals that of variables, there will be only one extremum center and the central value will be "1"). And if too few PCs are rotated, there will be too many extremum centers and the central values will be too homogeneous, thus the local features are hardly reflected. Here we use the "SHELF" method developed by North (1982) to choose the position of truncation points (Li and Tu, 2002).

3. Regionization for China's annual mean surface air temperatures

PCA analysis is performed on the annual mean temperature field from 1954 to 2002. According to the above methodology, the first 8 PCs (their cumulative variance contributions add up to 88%) are chosen to make RPCA, and the loadings of each RPC were calculated as the correlations between each RPC and every annual temperature anomaly series.

The results indicate that there is evident regional surface air temperature change in China.

Five regions are divided according to the loadings of PCs and their geographic distribution (those stations whose loading is above 0.5 and locate in adjacent areas are classified to the same region) over the whole country (not including data from Chinese HK, Chinese Macao and Chinese Taipei). The first region is the Yangtze River Basin (R1) and the second includes Northeast China and a part of North China (R2). The third region covers the whole of Southeast China (R3). The fourth region includes the west part of North China and Northwest China (R4). The fifth region includes the major part of Guangdong Province and south of Guangxi province and Hainan province (R5). These five regions include the major part of China (more than 95% of the 390 stations). These divisions are consistent with those given by Tu (2000) based on 160 stations, though there are some differences: R3 in our division includes 2 subregions in Tu's division, and R4 includes 3 subregions in Tu's division (Fig. 2). That is, the number of subregions in our division is less than that in Tu's division, and some of our subregions cover larger areas. This may be related to the differences of the dataset we adopted, the homogenization of the climatic data may introduce higher correlation between two time series, and denser data points make the single series less important during the RPCA. Given this, the above division is climatologically reasonable.

Further, we applied Principal Component Analysis (PCA) to the annual temperature anomaly series, and found that the variance contribution of the first PC of all the other subregions were above 70%, except that the variance contribution of the

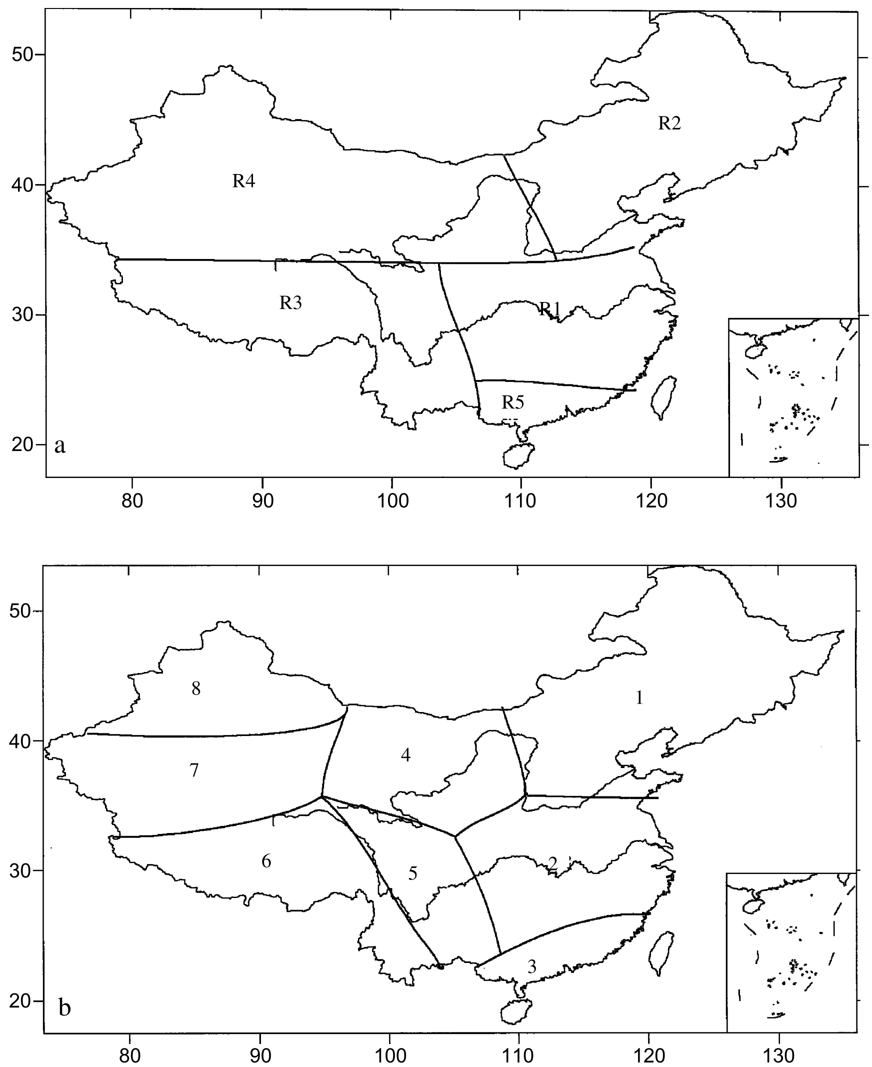

Fig. 2. Subregions for China's annual temperature variation (**a**) as compared with the Tu's division (**b**) (From Tu et al., 2000)

first PC of the third subregion (R3) is 65.4%. This indicates that all the anomaly changes in the same subregions are similar, and in every subregion, the first PC would reflect the main characters of the annual temperature change.

4. The statistical characteristics of annual mean temperature in cities with different populations in all subregions

Based on the above division, we analyze the magnitudes of the UHI effect on regional temperature change in each subregion. Easterling et al. (1997) took a population of $5 \cdot 10^4$ as the critical value of "urban" and "non-urban" stations. This means that in the "urban" stations, the annual temperature series should be affected by the UHI effects, while in the non-urban stations, UHI would have minor effects on the annual temperature change. Similar criteria were used in this study. Say, for a city with a population of more than $5 \cdot 10^4$, and whose station is located in the center of the city or for one with a population of more than $5 \cdot 10^5$ and whose observation is located in the center of the city in the suburb, the station should be regarded as an "obvious UHI station", while the rest are "non obvious urban stations". We obtained the results in Table 1.

For the three types of stations, i.e. stations in a city of more than $5 \cdot 10^5$ as type 1, population between $5 \cdot 10^4$ and $5 \cdot 10^5$ as type 2, and less than $5 \cdot 10^4$ as type 3. Statistics of the annual mean temperature in different subregions are listed in Table 2. These statistical characteristics (the average temperature values, the mean square deviations and the linear trends of the annual temperature change) of different types of stations in different subregions represent the relationships between the annual temperature and the population of the cities in each subregion.

Table 1. Numbers of "obvious UHI stations" in each subregion

Subregion	Total numbers in each subregions	Stations whose population ≥500,000	City station whose population ≥50,000	"Obvious UHI stations"
1	116	22	12	34
2	114	36	12	48
3	42	2	6	8
4	68	6	6	12
5	33	8	8	16
Sum	373	74	44	118

Trends are always used to describe long-term changes of climate variables because they can show us the magnitude of increase/decrease per year, decade or century. In this study, temperature is regarded as a predictand and the time is a predictor, then the linear regression is built:

$$\hat{y}_i = a + b\hat{x}_i \quad (i = 1, 2, \ldots, n) \tag{1}$$

The regression coefficients a and b are determined by least square procedure,

$$a = \bar{y} - b\bar{x} \tag{2}$$

$$b = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} \tag{3}$$

where $\bar{x}$, $\bar{y}$ present respectively the average value of predictand (temperature) and predictor (year), and b is the trend coefficient as the warming rate.

Shown in Table 2, all subregions, the statistics of annual mean temperature of different types of stations are different. For annual average temperature, the values of the "obvious UHI stations" (the first 2 types) are significantly higher

Table 2. Statistics of annual mean temperature of different types of stations in 5 subregions

	R1			R2			R3			R4			R5		
	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
Average (°C)	16.4	16.3	15.4	9.4	6.7	6.1	15.2	15.5	9.9	7.3	8.7	7.3	21.8	21.8	20.5
Mean square deviation (°C)	1.30	2.46	3.26	4.04	4.63	4.27	0.30	3.51	6.20	1.61	2.38	2.60	1.24	1.29	2.05
Linear trend (°C/decade)	0.136	0.089	0.063	0.319	0.340	0.346	0.145	0.151	0.163	0.335	0.305	0.282	0.182	0.153	0.151

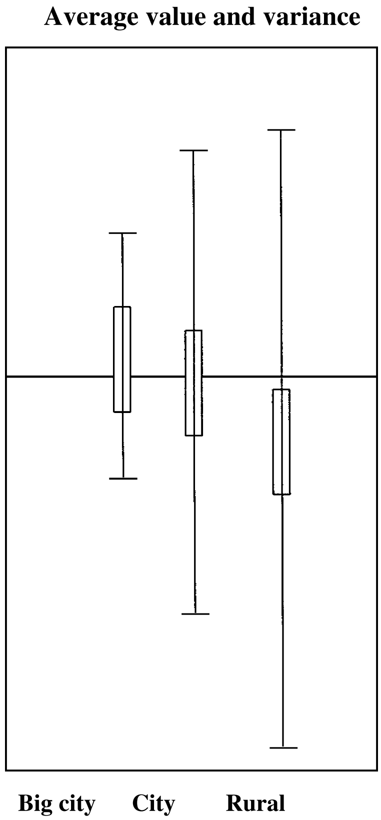

Fig. 3. The sketch map of mean square deviations of big cities, cities and rural stations' series

than those of the rest of the stations (the 3rd type). This may be due to two reasons. Firstly, it is due to the UHI effect (the urban site is always warmer than the suburb or rural site), and secondly, as in R3, many rural stations are located in some high-altitude regions, where it is colder than the lower-altitude regions. So the 3rd type (non UHI stations) is obviously colder than the 1st and 2nd types (obvious UHI stations). Former studies (Wang et al., 1990) quantified the UHI effect on the average temperature. As a matter of fact, quantification must be based on very detailed data analysis; otherwise, more or less biases will be induced. The other obvious feature is that the mean square deviations of the "obvious UHI stations" are much lower than those of the "non obvious UHI stations". This suggests that change ranges of annual temperature for the former type of station is less than those for the latter ones, due to the influence of urban pollution and other factors which reduced the diurnal range. As for the linear trends, these display some regional characteristics, especially in some regions such as the Yangtze River Basin, South China and Northwest China (R1, R4, R5), where the UHI effect enhanced the warming trends. In other regions such as Northeast and Southwest China (R2, R3), the UHI effect is opposite. These results are very important for research on UHI effects on regional climate change. They will be further discussed in this paper.

5. Average UHI effects on temperature series

In the previous sections we discussed that in different types of stations the UHI effect would be different. Therefore, the statistical characteristics of the annual mean temperature vary from station to station. The problem is how to assess the magnitude of the UHI effect in regional climate change. Since many factors such as topography and altitude may bring biases into the research results when using average temperatures, here we adopt temperature anomalies to denote the inter-annual variation of the regional temperature.

To generate regional anomaly series (Jiang, 2000), in each subregion, PCA was performed, and then the loading of the first PC was used as the weighting coefficient to calculate the weighted average temperature anomaly series. It is obvious that the weighted average series can rule out the effects of some single series with problems in the region, and can better reflect the regional temperature change signal.

Curves of five regional temperature anomalies are shown in Fig. 4 (here it is the curve without the "obviously UHI stations", which represents the background temperature change of these regions). From these pictures, we can see it is the universal warming trend in the whole of China, but to each subregion, the magnitudes of warming are different. The most significant warming region is Northeast China, and Northwest and North China the next one, then, Southwest, South China and the Yangtze River Basin. According to our calculation, during the last 50 years, the warming of R1's annual temperature is about 0.4 °C; R2 is the strongest warming region, where annual temperature is 1.6 °C warmer than 50 years ago. For R3, it is 0.7 °C,

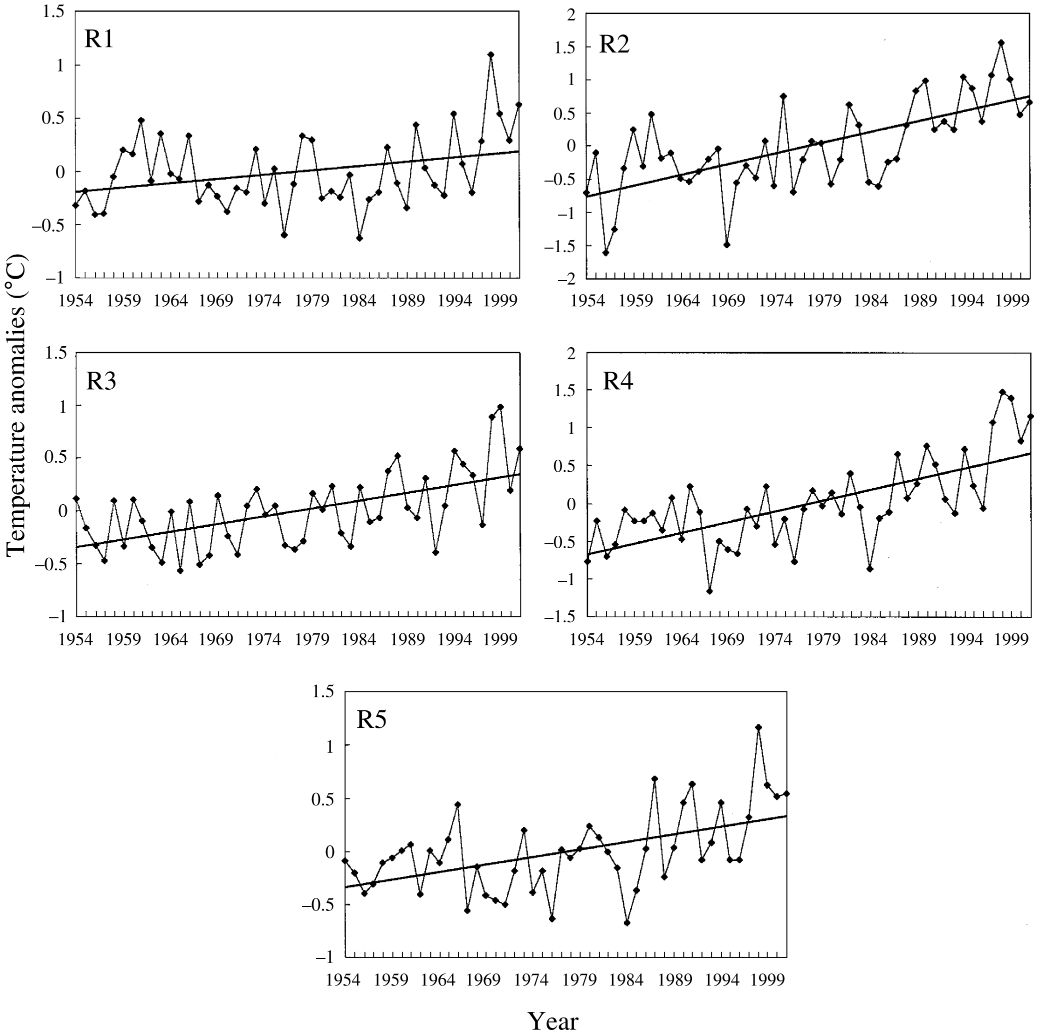

Fig. 4. Regional annual temperature anomalies in each subregions ("obvious UHI station" are excluded)

for R4 it is 1.4 °C and for R5 it is 0.7 °C. In addition, there are different features, even in the inter-decadal variation. In R1 region, annual temperature shows mainly the negative anomalies before the 1990s (especially in the late 1950s, late 1960s and early and mid 1980s), and from the 1990s on, there are mainly the positive anomalies. The warming has mainly occurred since the mid 1980s, and the warmest year during the last 50 years is 1998. Compared with the R1 region, in R2 region, the positive anomaly period was from the early 1980s; there are 3 very cold years, i.e. 1956, 1957 and 1969, while 1998 is also the warmest year. The inter-annual oscillation in R3 region is obvious: the warmer years and colder years happened by turn, even in the 1990s, 1993 and 1997 there are negative anomalies, and the warmest year is 1999, which is different to the former two subregions. The R4 region and R5 region are related to the same PC, so their change characters are similar, they both show positive anomalies since the mid 1980s, but they still have some slight differences in the linear trend of temperature. From all the above analysis we can see that the regional annual temperature change during the last 50 years is obvious.

In order to assess the impact of UHI effects on annual temperature, we calculated the regional annual temperature anomaly series with/without "obvious UHI stations" separately, then we compared the linear trends between each other in each subregion (Table 3). The differences between them are regarded as the average UHI effects on the annual temperature trends during the last 50 years.

Table 3. The regional temperature trends with/without "obvious UHI stations" during 1954–2001

Subregion	With UHI (°C/10a)	Without UHI (°C/10a)	Differences (°C/10a)	Average effects/ Background warming (°C)
1	0.094	0.083	0.011	0.055/0.42
2	0.324	0.326	−0.002	−0.010/1.63
3	0.142	0.148	−0.006	−0.030/0.74
4	0.288	0.286	0.002	0.010/1.43
5	0.149	0.138	0.011	0.055/0.69

Obviously, according to the above table, the warming trend including UHI stations in R1 is 0.093 °C/decade. After the "obvious UHI stations" excluded, it is 0.083 °C/decade. Assuming that the latter one represents the background change trend of R1 region, the average UHI effects of the R1 region should be 0.011 °C/decade, and the rest may be deduced by analogy. The average UHI effects in different subregions should be quantified, for example, in the Yangtze River Basin and South China (R1, R5), the overall warming by UHI effect during 50 years is about 0.55 °C. In the other 3 subregions (including North west, North China, Northeast and Southwest China), however, the UHI effects are about 0.01 °C, −0.01 °C and −0.03 °C. When compared with the background trends of these subregions, they all very small. So we conclude that the UHI effects in China during the last 50 years are minor, which also supports Peterson's (2003) results for UHI effects in the US. Generally, the maximum UHI effect is less than 0.06 °C in the last 50 years. For subregions 2 and 3, the differences are negative (Table 3). Is it only because of the numerical calculation error? Or is it an observation fact? We found from Table 2 that in these two subregions, the trends are decreasing with increasing population, which is obviously different from the other subregions (the subregions 1, 4 and 5). It is not difficult to understand the negative differences between the trends with/without the obvious urban stations, though it is very difficult to interpret its physical causes.

Based on a single station, Seoul, which is not far from China in Korea, Kim and Baik (2002) found an UHI effect of about 0.56 °C during the last 24 years. It is much larger than our results. This suggests that to study the UHI effects of a local urban station is a complicated task. Denser observation data (including satellite, radar and special observation data) and more detailed analysis on local scales may help to answer this remaining question.

6. Discussion and conclusions

In this study, based on homogenized temperature data and climate regionization for annual mean temperature series in China, the classification of annual temperature is performed, and the regional annual temperature variation is analyzed. The UHI effects on all subregions are discussed. The following are the main results:

1) The low frequency oscillation of annual temperature in China shows obvious regional features. From 1954 to 2001, the annual mean temperature field can be divided into five larger regions. To test the rationality of this division, PCA was performed on annual temperature series in each subregion, and it is found that they have similar climate change. Based on that conclusion, the background warming trend during the last 50 years is revealed: North east, North west and North China are the regions with most significant warming with temperature increases over 1.5 °C. Southwest and South China have warmed about 0.7 °C. The least warming region is the Yangtze River Basin, where the temperature rise is less than 0.5 °C.
2) No matter in what subregions, the statistics of the annual mean temperature of different types of stations are different (Fig. 3). For annual mean temperature, although the UHI effects have been quantified in this paper, the average values of the "obvious UHI station" are significantly higher than those of the other stations; the other outstanding feature is that the mean square deviations of the "obvious UHI stations" are much lower than those of "non obvious urban stations". This shows that change in the former types of station annual mean temperatures is less than change in the latter station types. This can be attributed to the urban pollution and other

factors which caused the redution in diurnal range.

3) There are warming trends reflected by annual mean temperatures in all the subregions. But for each subregion, the magnitude of warming is diverse, and the UHI effects are also different. In South China and the Yangtze River Basin, the UHI effect has enhanced warming more than that in North China and Northwest China. In the Southwest and Northeast, the effect is the reverse, which deserves more study.
4) In all subregions, the UHI effect in China during the last 50 years is minor compared to the background trend of increasing temperature.

Acknowledgements

The authors would thank the following persons: Dr. Tom Peterson, who shared his research results with the authors unselfishly and gave some good suggestions; Prof. Tu Qipu, Prof. Jiang Zhihong, who gave much helpful discussion during this study. Especially we thank Prof. Zhai Panmao for his kind help of revising the language of the paper. The research was supported by National Meteorological Center of China Meteorological Administration, and partly by Jiangsu Province Key Laboratory of Meteorological Disaster and Environmental Variation Opening Task (No. KJS02104).

Appendix: some methodologies referred in this paper:

Easterling-Peterson techniques for homogenization of the climatological data

Solow (1987) described a technique for detecting a change in the trend of a time series by identifying the change point in a two-phase regression where the regression lines before and after the year being tested were constrained to meet at that point. Since changes in instruments can cause step changes, Easterling and Peterson (1995) developed a variation on the two-phase regression in which the regression lines were not constrained to meet and where a linear regression is fitted to the part of the (candidate-reference) difference series before the year being tested and another after the year being tested. This test is repeated for all years of the time series (with a minimum of 5 years in each section), and the year with the lowest residual sum of the squares is considered the year of a potential discontinuity. A residual sum of the squares from a single regression through the entire time series is also calculated. The significance of the two-phase fit is tested with (i) a likelihood ratio statistic using the two residual sums of the squares and (ii) the difference in the means of the difference series before and after the discontinuity as evaluated by the Student's *t*-test. If the discontinuity is determined to be significant, the time series is subdivided into two at that year. Each of these smaller sections are similarly tested. This subdividing process continues until no significant discontinuities are found or the time series are too short to test (shorter than 10 years). Each of the discontinuities that have been identified are further tested using a multi-response permutation procedure (MRPP; Mielke, 1991). The MRPP test is non-parametric and compares the Euclidean distances between members within each group with the distances between all members from both groups, to return a probability that two groups more different could occur by random chance alone. The two groups are the 12-year windows on either side of the discontinuity, though the window is truncated at a second potential discontinuity. If the discontinuity is significant at the 95% level (probability $(P) = 0.05$), it is considered a true discontinuity. The adjustment that is applied to all data points prior to the discontinuity is the difference in the means of the (station-reference) difference series' two windows.

PCA

Principal components analysis (PCA) is used to explain the total variability of p correlated variables through the use of p orthogonal principal components. The components themselves are merely weighted linear combinations of the original variables. The first principal component can be expressed as follows,

$$Y1 = a11X1 + a21X2 + \cdots + ap1Xp \tag{A1}$$

or in matrix form:

$$Y1 = a'x \tag{A2}$$

The aj1 are scaled such that $a1'a1 = 1$. Y1 accounts for the maximum variability of the p variables of any linear combination. The variance of Y1 is $\lambda 1$.

Next, principal component Y2 is formed such that its variance, $\lambda 2$ is the maximum amount of the remaining variance and that it is orthogonal to the first principal component. That is, $a1'a2 = 0$.

One continues to extract components until some stopping criteria is encountered or until p components are formed. It is possible to compute principal components from either the covariance matrix or correlation matrix of the p variables. If the variables are scaled in a similar manner than many researchers prefer to use the covariance matrix. When the variables are scaled very different from one another than using the correlation matrix is preferred.

The weights used to create the principal components are the eigenvectors of the characteristic equation,

$$\begin{aligned}(S - \lambda iI)a &= 0, \quad \text{or}\\ (R - \lambda iI)a &= 0\end{aligned} \tag{A3}$$

where S is the covariance matrix and R is the correlation matrix. The λi are the eigenvalues, the variances of the components.

PCA is a great tool for reducing the variability in the data into a few PCs, in this paper the first PC interpreted at least 65% of the total variance of the subregional temperature field, so the loading of the first PC is used as the weighing of the each single station to create the regional temperature anomalies.

RPCA

The primary advantage of the PCA is its ability to compress the complicated variability of the original data set into a relatively few temporally uncorrelated components. The temperature field ${}_pX_n$ is decomposed to the loading vector matrix ${}_pL_k$ and PC matrix ${}_kF_n$,

$$ {}_pX_n = {}_pL_k\,{}_kF_n \tag{A4}$$

However, it has certain characteristics, that is: domain shape dependence, subdomain instability, sampling problems and inaccurate portrayal relationships embedded within the input data, which severely limit its usefulness in certain situations. In this paper, a rotated principal components solution refers to a linear transformation (rigid rotation) of the initial principal components utilizing the varimax method (Horel, 1981). The varimax method maximizes the variance of the squared correlation coefficients between each rotated principal component and each of the original time series. That is,

$$ \mathrm{B} = \mathrm{L}\Gamma, \quad \mathrm{G} = \Gamma'\mathrm{F} \tag{A5}$$

Γ is an orthogonal matrix, $\Gamma\Gamma' = 1$, B is rotated loading vector (RLV) and G is rotated PC (RPC). The detailed arithmetic of the RPCA and the varimax method is given by Kaiser (1958).

References

Easterling DR et al (1997) Maximum and minimum temperature trends for the global. Science 277: 364–367

Gallo KP, Easterling DR, Peterson TC (1996) The influence of land use/land cover on climatological values of the diurnal temperature range. J Clim 9(11): 2941–2944

Hansen JE, Ruedy R, Sato Mki, Imhoff M, Lawrence W, Easterling D, Peterson T, Karl T (2001) A closer look at United States and global surface temperature change. J Geophys Res 106: 23947–23963

Horel JD (1981) A rotated principal component analysis of the interannual variability of the Northern Hemisphere 500 mb height field. Mon Wea Rev 109: 2080–2092

Huang J (2000) Statistic analysis and forecast methods in meteorology, 2000. Beijing: Meteorological Press, 126–134 pp

IPCC (2001) Climate change 2001: The Science of Climate Change. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton JT, Ding Y etc.]

Jiang Z (2000) Study on temporal and spatial structures of the low/high frequency variability for global temperature fields with their interaction and causes. Doctoral dissertation of Nanjing Institute of Meteorology, June

Kaiser HF (1958) The varimax criterion for analytic rotation in factor analysis. Psychometrika 23: 187–200

Kalnay E, Cai M (2003) Impact of urbanization and land-use change on climate. Nature 423: 528–531

Karl TR, Diaz HF, Kukla G (1988) Urbanization: its detection and effect in the United States climate record. J Clim 1: 1099–1123

Kim Y, Baik J (2002) Maximum Urban Heat Island Intensity in Seoul. J Appl Meteorol 41: 651–659

Li Q, Liu X, Zhang H et al (2004) Detecting and adjusting on temporal inhomogeneity in Chinese Mean Surface Air Temperature Dataset. Adv Atmos Sci 21(2) (in press)

Li Q, Tu Q (2002) The regional characters of annual precipitation in northern hemisphere land-surface and China and its correlation. Journal of Nanjing Institute of Meteorology 26: 92–99

Liu X, Li Q (2003) Research of the inhomogeneity test of climate data series in China. Acta Meteorological Sinica 17: 492–502

Mielke PW (1991) The application of multivariate permutation methods based on distance functions in the earth sciences. Earth-Sci Rev 31: 55–71

Murry Mitchell J Jr (1961) The temperature of cities. Weatherwise 12: 225–229

North GR, Bell TL, Cahalan RF, Moeny FJ (1982) Sampling errors in estimation of empirical orthogonal functions (EOF). Mon Wea Rev 110: 699–706

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

Peterson TC, Easterling DR et al (1998) Homogeneity adjustments of in situ atmospheric climate data: a review. Int J Climatol 18: 1493–1517

Peterson TC (2003) Assessment of urban versus rural in situ surface temperature in the contiguous United States: No difference found. J Clim 16: 2941–2959

Portman DA (1993) Identifying and correcting urban bias in regional time series: surface temperature in China's northern plains. J Clim 6: 2298–2308

Solow A (1987) Testing for climatic change: An application of the two-phase regression model. J Climate Appl Meteor 26: 1401–1405

Tu Q, Deng Z, et al (2000) Studies on the regional characters of air temperature abnormal in China. Acta Meteorological Sinica 58: 289–296

Wang Wei-Chyung, Zhaomei Zeng, Karl TR (1990) Urban Heat Islands in China. Geophy Res Lett 17: 2377–2380

Zhai P, Eskridge RE (1996) Analysis of inhomogeneities in radiosonde and humidity time series. J Clim 9: 884–894

Authors' addresses: Q. Li (e-mail: liqx@cma.gov.cn), H. Zhang and X. Liu, Division of Meteorological Data, National Meteorological Center, China Meteorological Administration, No. 46, ZhongGuanCun, South avenue, Haidian District, Beijing, China 100081; J. Huang, Physical Institute, Peking University, Beijing, China 100872.