Supporting Text

Chinese Meteorological Station Network

The Chinese observed land-surface air temperature data set includes measurements from 731 meteorological stations from 1951 to the present, as collected and processed by the National Meteorological Center of the China Meteorological Administration (1, 2). Because some stations were removed, the actual total number of stations in operation today is 671, distributed among 31 provinces and municipalities. Fig. 6 shows the district map of these provinces and municipalities, and Fig. 7 shows the location and topography of these stations. In this study, we use the monthly mean daily maximum and minimum temperature data (1) from the 610 stations that have a complete set of observations for the period from January 1979 to December 1998.

China, with a population of ≈1.3 billion and an area of 9.6×10^6 km^2, has a complex topography. Its terrain descends gradually from west to east like a staircase, with the towering Tibetan plateau called the "roof of the world" to the west and the flat and fertile plains to the eastern coast of the Pacific Ocean. From north to south, the elevation drops from 1,000–2,000 m of the Inner Mongolia Plateau to <200 m of southeast China. This varied topography is associated with a large gradient in climate. China has a marked continental monsoon climate, with cold and dry winters and hot and humid summers, especially in southeast China. Northerly winds prevail in winter, whereas southerly winds reign in summer. The warm and moist summer monsoons from the oceans bring abundant rainfall and high temperatures to most of China. Annual precipitation varies greatly from <50 mm in Northwest China to ≈3,000 mm in southern China (1).

Problems in Estimating UHI Effects by Using Observational Data

The UHI effect has been estimated by using *in situ* observations around the world, mostly by comparing observed temperatures in urban stations with those from surrounding rural stations (3, 4). The estimated UHI varies significantly by region, time, and method. In the

U.S., the estimated UHI varies from 0.06°C to 0.15°C per century, depending on whether population data or satellite measurements of night lights are used to classify urban versus rural stations (5, 6). In contrast, Peterson (4) finds that the UHI has no significant impact on temperature in the U.S. after observed temperatures are adjusted for inhomogeneities caused by "nonclimatic" factors such as changes in location (altitude and latitude), observation time, instrumentation, and nonstandard siting. These nonclimatic factors could introduce artifacts in long-term observations and rural–urban differences and thus may bias the estimate of UHI. The lack of an UHI effect may be caused by micro- and local-scale impacts overwhelming the mesoscale UHI. Industrial sections of towns may well be significantly warmer than rural sites, but urban meteorological observations are more likely to be made within cool "park islands" than industrial regions (4). Evidently, the UHI is more complex than usually considered.

Using temperature differences between urban and rural stations to estimate the UHI effect in China may be inappropriate (7) for several reasons. First, most Chinese stations are located in or near cities, with only a few in mountainous or remote regions or on small islands. For example, the China's Fifth (2000) Census indicates that only 27% of Chinese meteorological stations have a permanent urban population <10,000, and these stations are located mostly in west China [Q. Li (China's National Meteorological Center) personal communication]. Although China is comparable in size to the U.S., it has considerably fewer meteorological stations and each city generally has only one station. For example, China's two biggest cities, Beijing and Shanghai, each has only one station available in the Chinese network. Kalnay and Cai (8) used nearly 2,000 meteorological stations for the continental U.S. in their study. It is impossible to find a corresponding rural station for most of the urban ones, especially in eastern and southern China. Consequently, if using the rural–urban difference to estimate the UHI, one possibly is comparing temperature between two different urban stations at regional scales or between two different regions at large scales. For this reason, Li *et al*. (7) divided China into five subregions and estimated their UHI effect separately. Furthermore, adjusting spatial and temporal homogeneities for *in situ* observations in China may inadvertently sacrifice the UHI effect because the adjustments often are performed by comparing a target station with its neighbors that generally are also urban stations and relatively far away. Second, China's rapid urbanization in the past two decades could transfer a station from rural into urban in a very short period. The continuous expansion in urban population and area makes the classification of urban versus rural station dynamic. Third, Chinese cities have a much higher density of population and urban buildings than do cities in most developed countries. Cities in the U.S. extend many kilometers to suburban areas where people reside and that can have as much vegetation as rural areas, whereas Chinese cities have a significantly higher density of population, residential buildings, shopping malls, schools, roads, etc., and much less vegetation than their neighboring rural areas do because people live within cities. These unique characteristics could make the UHI effect more pronounced in China than in other countries like the U.S.. For example, after the homogeneity adjustments to China's mean surface temperature, Li *et al*. (7) find that most Chinese temperature time series are inevitably affected by UHI.

Homogeneity Assessment of Observational Data

The central problem with any long-term analysis of climate data is that inhomogeneities, which are caused by several factors such as changes in location, observing practices, instrumentation, and nonstandard siting (4, 9, 10), could introduce large biases in the data and thus lead to inaccurate or erroneous conclusions. Several techniques have been introduced to remove these factors (9). They generally compare a reference series against a candidate time series to test for inhomogeneities. The reference series is created by using neighbor stations to establish an ideal, completely homogenous series. The candidate series then is adjusted by comparison with this reference series. Currently, the Global Historical Climatology Network (GHCN) (11) and the U.S. Historical Climatology Network (USHCN) (12) are two homogeneity-adjusted time series at large scale for long-term climate analysis. However, the GHCN data set could not be used in this study because most of the data between 1979 and 1998 are not available over China.

Recently, Li *et al.* (2) adopted the Easterling–Peterson (E-P) techniques (12) to test Chinese meteorological observations for inhomogeneities in historical mean surface air temperature series from 1951 to 2001. The results indicate that the time series have been affected greatly by inhomogeneities due to the station relocation and other nonclimatic effects. Based on the amplitude of changes in the first difference of the time series and the monthly distribution features of surface air temperatures, discontinuities identified by applying the E-P technique supported by China's metadata, or by comparison with other approaches, have been adjusted. The inhomogeneity testing detects most nonclimatic changes and indicates that the adjusted data has been largely improved in its reliability and could help decrease uncertainties in the study of observed climate change in China.

Here we cannot make the same adjustments to the maximum and minimum temperature time series used in this study because the China's National Meteorological Center does not allow foreign scientists access to the required metadata. Instead, we use the latest homogeneity-adjusted mean surface temperature data set of Li *et al.* (2) to assess the magnitude of nonclimatic effects in China. Fig. 8 shows the total number of main discontinuities in the annual mean air temperature for each station during the period 1951–2001 due to station relocations and other nonclimatic effects. Most discontinuities are located in north and west China, with only fewer in southeast China, and some stations have up to six discontinuities. Fig. 9 illustrates the long-term trends of annual mean temperature before and after the adjustment. Significant adjustments are observed in Qinghai, North China, Tibet, and Sichuan. Evidently, the homogeneity adjustments are minimal in southeast China.

Quality Assessment of R-2 Data

Kalnay and Cai (8) estimate the impacts of urbanization and other land-use changes on climate based on the difference in surface temperature trends between meteorological observations at 1,982 surface stations in the continental U.S. and NCEP/NCAR Reanalysis (R-1) (13). The R-1 uses the most extensive observations available from a variety of sources including ship, rawinsonde, pibal, aircraft, and satellite, etc., to assimilate these data with an assimilation system kept unchanged. The R-1 data are strongly influenced by atmospheric vertical soundings of wind and temperature, and surface temperatures are estimated from the atmospheric values (surface observations of temperature, moisture, and wind over land are not used) and thus are insensitive to changes in land surface (8). Therefore, the differences in surface temperature trends between meteorological observations and R-1 are postulated to represent the impacts of urbanization and other land-use changes on climate.

This method assumes that the quality of R-1 surface air temperatures is satisfactory. One known deficiency with R-1 data is its poor performance in the description of cloudiness and surface moisture, which could bias the computation of the surface energy budget and thus surface air temperature (14, 15). Increased cloud cover is linked with the worldwide decline in DTR, and increased soil moisture could reduce DTR through enhanced evapotranspiration (14–16). Consequently, differences in clouds and soil moisture between observed and R-1 data could contaminate the UHI estimate. The second deficiency with R-1 data is its poor performance over mountainous regions (8). The model of R-1 has a spatial resolution of 2.5° and thus uses a land-surface boundary that is smoother than reality. This smoothing could introduce large biases in the model's altitude or land-surface properties relative to the actual meteorological stations and thus in the R-1 temperatures over mountainous areas with varied topography. Vose *et al.* (10) and Peterson (4) point out the lack of inhomogeneity adjustments in the observational data in the study of Kalnay and Cai (8). Trenberth (15) argues that the R-1 does not include effects of changing atmospheric composition such as greenhouse gases and aerosols on radiative forcing, but Cai and Kalnay (17) have shown in their reply that the R-1 data are able to capture the full strength of climate trends because the reanalysis assimilates atmospheric temperatures and other observations that are affected by the greenhouse gases and aerosols.

Here we adopt the method of Kalnay and Cai (8) to estimate the impact of urbanization and other land-use changes on climate in China but pay more attention to the aforementioned problems. We choose the NCEP/DOE AMIP-II Reanalysis (R-2) (14) covering 1979–present at spatial resolution of ≈1.9° instead of R-1. Although based on the widely used R-1, the R-2 has improved its quality by featuring newer physics and observed soil moisture forcing and also by fixing known errors of R-1. For example, the soil wetness evolution is treated completely differently in R-2 than in R-1, and a new cloudiness-relative humidity table is generated to fix the errors in R-1. Consequently, the R-2 should more accurately characterize soil moisture, cloud, and near surface temperature over land (14).

One way to evaluate the accuracy of R-2 data is to compare the time series of monthly maximum and minimum temperature anomalies with observed data. We calculate the correlation coefficient between the two data sets for both maximum and minimum temperatures for all stations. If the R-2 captures well the observed surface temperature variations due to changes in weather systems, they should be highly correlated. Fig. 10 shows the spatial pattern of correlation coefficients between the R-2 and observed time series of maximum and minimum temperatures. Evidently, the correlation coefficients are greatest for southeast China, followed by north China, whereas west China has the smallest correlation coefficients. This pattern corresponds with China’s topography. The smallest correlation is observed at some stations in Tibet, Sichuan, and Yunnan provinces, where the topography is highly variable. Similar results also are found for several stations located on high mountains or islands in east China. These results suggest that the quality of the R-2 data need to be checked before it can be used in climate studies.

Choosing Our Study Region

To ensure the reliability of our analyses, we choose our study region carefully based on the quality, reliability, and homogeneity of the observational and R-2 data as described above. Our study region should have the smallest nonclimatic effects in the observations and the highest correlation coefficients between the observed and R-2 data. It also should include the area where most of China’s urbanization has occurred. China has experienced a slow urbanization due to its special political, social, and economic circumstances before its reforms in 1978 (18). Since then, its rapid urbanization has been very inhomogeneous and occurred mainly in southern and eastern provinces, with the fastest economic growth near the Yangtze and Pearl River deltas (19).

Evidently, these requirements are satisfied by southeast China. This region includes twelve provinces (Anhui, Guangdong, Guangxi, Jiangsu, Jiangxi, Henan, Zhejiang, Hubei, Hunan, Fujian, Guizhou, and Hainan) and two municipalities (Shanghai and Chongqing). We eliminated the Hainan province from this study, which consists of islands surrounded by oceans, and several stations located on mountains and small islands in other provinces for two reasons. One is the difference in altitude and land-surface properties between the coarse resolution R-2 data and the observations. The second is that the size of some islands is smaller than that of NDVI pixels (64 km^2), which have observations only over land. Consequently, we focus our study on southeast China (20°N–36°N, 102°E–123°E), consisting of 194 spatially well distributed stations and representing an area where most of China's urbanization has occurred (18). This region has (*i*) the highest meteorological station density; (*ii*) the most uniform station distribution; (*iii*) the minimal nonclimatic effects; and (*iv*) the best consistency between the observed and R-2 data in China.

Following Kalnay and Cai (8), we also test the sensitivity of the R-2 data to urbanization by comparing the annual mean temperature trends between urban and rural stations classified based on population data. Because only cities with populations larger than 100,000 are available from the China's Fourth (1990) Census (United Nations, *Population of Capital Cities and Cities of 100,000 and More Inhabitants: China*, available online from http://unstats.un.org/unsd/citydata/default.asp?cid = 157), we adopt a threshold of 100,000 instead of 50,000 or less, a criterion that is often used to differentiate between an urban or nonurban station (5, 6, 20, 21), to classify all 194 stations into two categories: 109 (rural) and 85 (urban). The annual mean temperature trend (in °C per decade) and its standard deviation are 0.27 ± 0.23 for rural stations and 0.30 ± 0.21 for urban stations in the R-2 data. In the observations, the corresponding values are 0.32 ± 0.21 and 0.39 ± 0.21, respectively. The rural–urban difference is statistically significant at the 5% level for the observations (0.07) but insignificant for the R-2 (0.03), indicating small sensitivity of the latter to UHI.

As discussed previously, using the rural–urban temperature difference may be inappropriate to estimate the UHI in China. For example, the above estimated rural–urban difference in the observational data may not represent the UHI effect because the majority of urban stations classified above are located in the coastal provinces, whereas most of the rural stations are located in the northwestern provinces. In other words, it may represent the temperature differences between the two regions rather than the differences between urban stations and their rural counterparts.

Seasonal Variations in Observed and R-2 Temperature Trends

Although the R-2 data are based on a better description of cloudiness and soil wetness, some uncertainties may still remain regarding the complexity of China's topography and climate. We examine the consistency of R-2 and observational data over our study region by season: winter (December–February), spring (March–May), summer (June–August), and autumn (September–November). Fig. 11 shows the histogram for the correlation coefficient between the R-2 and observed data for both maximum and minimum temperatures by season. Evidently, the R-2 shows the best consistency with the observational data during winter, followed by autumn, spring, and summer. The weakest consistency in summer indicates that the R-2 temperatures may be still biased by its incomplete cloud and soil moisture description. These results imply that the winter R-2 data will generate the most reliable estimate for the UHI effect.

The UHI effect in China is largest during spring based on analyses of *in situ* observations by using the rural–urban differences [Q. Li, (China's National Meteorological Center), personal communication], but spring corresponds to the largest aerosol effect (22). Therefore, we use the winter temperature data to estimate the urbanization effect on climate in China. Winter is also the season when the effects of clouds and soil wetness are smallest for observational data. Increased cloud cover has been linked with the worldwide decline in DTR (16). Urban and rural areas may differ in cloud cover and rainfall, and this difference should be largest in summer, especially in a marked monsoon climate country like China. Therefore, the UHI should be expected to be more visible in winter than in summer when both clouds/rainfall and UHI decrease DTR and thus cannot be differentiated in the observations.

To estimate the magnitude of the nonclimatic effects over our study region, we use the original and homogeneity-adjusted annual mean temperature data of Li *et al.* (2) to compute the difference in temperature trends for the period of 1979–1998. Our estimate is 0.002°C per decade, indicating the average nonclimatic effect over our study region is small although the homogeneity adjustments could be large for a specific station. This small effect may be, in part, attributed to two factors: (*i*) calibrations are made to most Chinese stations when they are relocated; and (*ii*) the relocations will not produce large differences in altitude and thus in temperatures due to the small variation in elevation over our study region. These results suggest that using the unadjusted data in southeast China will not introduce significant biases.

We also calculate temperature trends and their relation with changes in percentage urban and NDVI for other seasons. Figs. 12–15 show the observed minus R-2 trends for maximum, minimum, and DTR for spring, summer, autumn and annual mean. Table 1 lists seasonal and annual mean temperature trends for the observations, R-2, and their differences. Evidently, the R-2 data for mean temperature has a much smaller trend during summer than in other seasons, resulting in a significant observed minus R-2 mean temperature trend. Table 2 lists the correlation coefficients between changes in percentage urban and observed minus R-2 trends in seasonal and annual DTR trends during the period of 1979–1998. The urban index (i.e., percentage urban) shows the largest correlation with changes in the DTR in winter, followed by autumn, spring, and summer. This ranking is consistent with that in the R-2 data quality in Fig. 11. Table 3 lists the correlation coefficients between summer NDVI trends and observed minus R-2 trends in seasonal and annual mean temperatures during the period of 1979–1998.

Variations in NDVI show the highest correlation with the minimum temperature, as reported in Gallo and Owen (23).

1. Fang, J. Y., Piao, S. L., Tang, Z. Y., Peng, C. H. & Ji, W. (2001) *Science* **293,** 1723.

2. Li, Q., Liu, X., Zhang, H., Peterson, T. C. & Easterling, D. R. (2004) *Adv. Atmos. Sci.* **21,** 1-9.

3. Arnfield, A. J. (2003) *Int. J. Climatol.* **23,** 1-26.

4. Peterson, T. C. (2003) *J. Clim.* **16,** 2941-2959.

5. Hansen, J., Ruedy, R., Sato, Mki., Imhoff, M., Lawrence, W., Easterling, D., Peterson, T. & Karl, T. (2001) *J. Geophys. Res.* **106,** 23947-23963.

6. Gallo, K. P., Owen, T. W., Easterling, D. R. & Jamason, P. F. (1999) *J. Clim.* **12,** 1344-1348.

7. Li, Q., Zhang, H., Liu, X. & Huang, J. (2004) *Theor. Appl. Climatol.*, in press.

8. Kalnay, E. & Cai, M. (2003) *Nature* **423,** 528-531.

9. Peterson, T. C., Easterling, D., Karl, T., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., *et al*. (1998) *Int. J. Climatol.* **18,** 1493-1517.

10. Vose, R. S., Karl, T. R., Easterling, D. R., Williams, C. N. & Menne, M. J. (2004) *Nature* **427,** 214.

11. Peterson, T. C. & Vose, R. S. (1997) *Bull. Am. Meteorol. Soc.* **78,** 2837-2849.

12. Easterling, D. R. & Peterson, T. C. (1995) *Int. J. Climatol.* **15,** 369-377.

13. Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., *et al*. (2001) *Bull. Am. Meteorol. Soc.* **82,** 247-268.

14. Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M. & Potter, G. L. (2002) *Bull. Am. Meteorol. Soc.* **83,** 1631-1643.

15. Trenberth, K. E. (2004) *Nature* **427,** 213.

16. Dai, A., Genio, A. D. Del & Fung, I. Y. (1997) *Nature* **386**, 665-666.

17. Cai, M. & Kalnay, E. (2004) *Nature* **427,** 214.

18. Liu, S., Li, X. & Zhang, M. (2003) in *Scenario Analysis on Urbanization and Rural–Urban Migration in China*, Interim Report IR-03-036 (International Inst. for Applied Systems Analysis, Vienna, Austria), www.iiasa.ac.at/Publications/Documents/IR-03-036.pdf.

19. Seto, K. C., Kaufmann, R. K. & Woodcock, C. E. (2000) *Nature* **406,** 121.

20. Easterling, D. R., Horton, B., Jones, P. D., Peterson, T. C., Karl, T. R., Parker, D. E., Salinger, M. J., Razuvayev, V., Plummer, N., Jamason, P., *et al*. (1997) *Science* **277,** 364-367.

21. Karl, T. R., Jones, P. D., Knight, R. W., Kukla, G., Plummer, N., Razuvayev, V., Gallo, K. P., Lindseay, J., Charlson, R. J. & Peterson, T. C. (1993) *Bull. Am. Meteorol. Soc.* **74,** 1007-1024.

22. Luo, Y., Lu, D., He, Q., Li, W. & Zhou, X. (2000) *Chinese Sci. Bull.* **45,**1328-1333.

23. Gallo, K. P. & Owen, T. K. (1999) *J. Appl. Meteor.* **36,** 806-813.

Table S1. Regional average seasonal and annual mean temperature trends (°C per decade) for the observations, R-2 and their differences during the period of 1979-1998 in southeast China.

		winter	spring	summer	autumn	annual
Mean	Observed	0.4498	0.4022	0.1397	0.4024	0.3485
	R-2	0.3998	0.3262	0.0200	0.3889	0.2837
	Difference	0.0500	0.0760	0.1196	0.0134	0.0648
Maximum	Observed	0.3522	0.3829	0.0898	0.5073	0.3330
	R-2	0.3681	0.3518	-0.0517	0.5583	0.3065
	Difference	-0.0158	0.0310	0.1416	-0.0509	0.0264
Minimum	Observed	0.5475	0.4216	0.1896	0.2975	0.3640
	R-2	0.4316	0.3006	0.0918	0.2196	0.2609
	Difference	0.1158	0.1209	0.0977	0.0778	0.1031
DTR	Observed	-0.1952	-0.0386	-0.0998	0.2098	-0.0309
	R-2	-0.0634	0.0511	-0.1436	0.3386	0.0456
	Difference	-0.1316	-0.0898	0.0438	-0.1288	-0.0766

Table S2. Correlation between changes in percent urban and observed minus R-2 trends in seasonal and annual DTR trends during the period of 1979-1998 in southeast China.

Season	winter	spring	summer	autumn	annual
Correlation coefficient	**-0.77**	0.13	0.03	-0.34	-0.26

Note: Coefficients in bold are statistically significant at $p<0.01$.

Table S3. Correlation between summer NDVI trends and observed minus R-2 trends in seasonal and annual mean temperature during the period of 1979-1998 in southeast China.

	winter	spring	summer	autumn	annual
Maximum	-0.51	**-0.71**	-0.12	*-0.58*	*-0.63*
Minimum	*-0.67*	**-0.77**	-0.14	*-0.61*	**-0.75**
DTR	0.53	-0.20	-0.01	0.06	0.09

Note: Coefficients in bold (italic) are statistically significant at $p<0.01$ ($p<0.05$).

Fig. S1. District map of China’s 31 provinces and municipalities (source: http://china.scmp.com/map/).

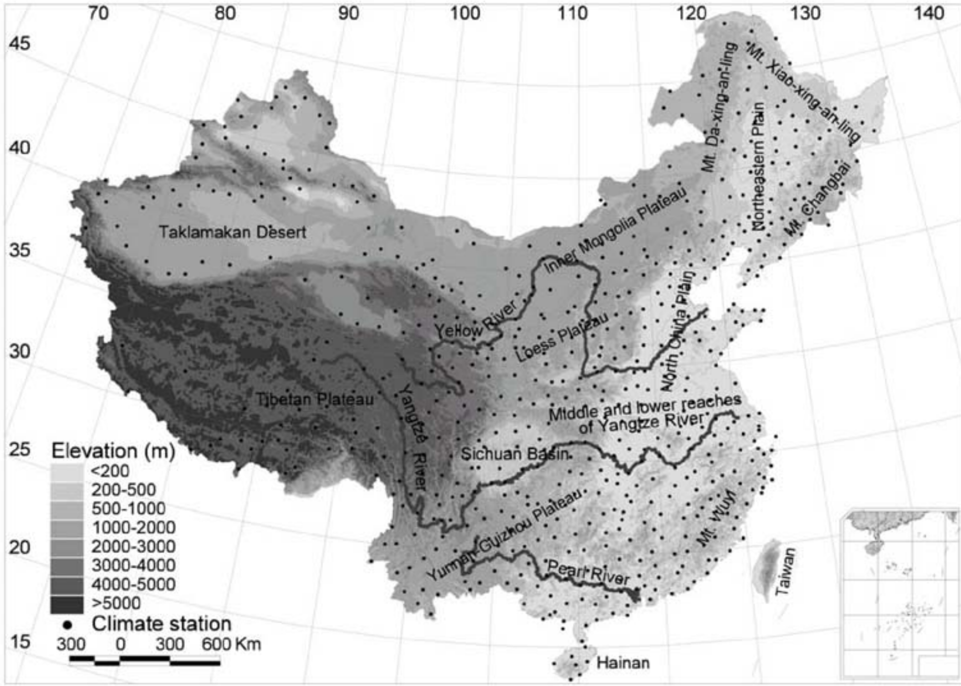

Fig. S2. Location and topography of Chinese meteorological station network.

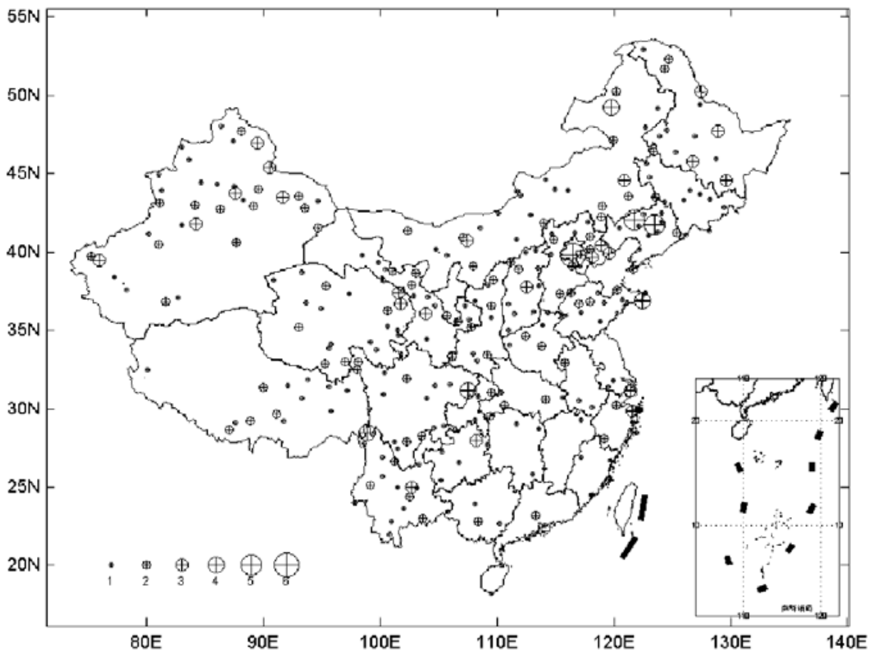

Fig. S3. Total number of main discontinuities for each station during the period of 1951-2001 (2).

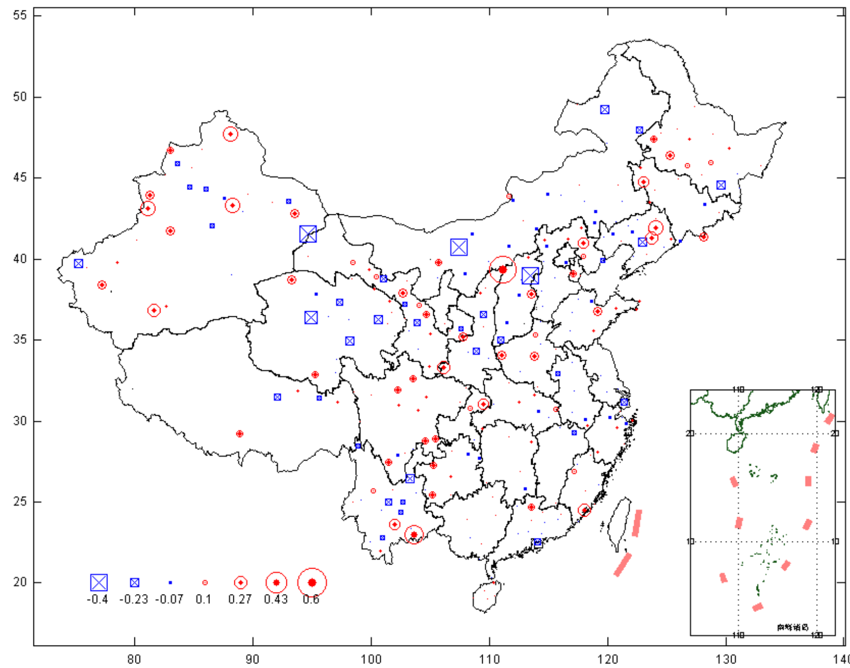

Fig. S4. Differences in annual mean temperature trend before and after the homogeneity adjustments during the period of 1951-2001 (2).

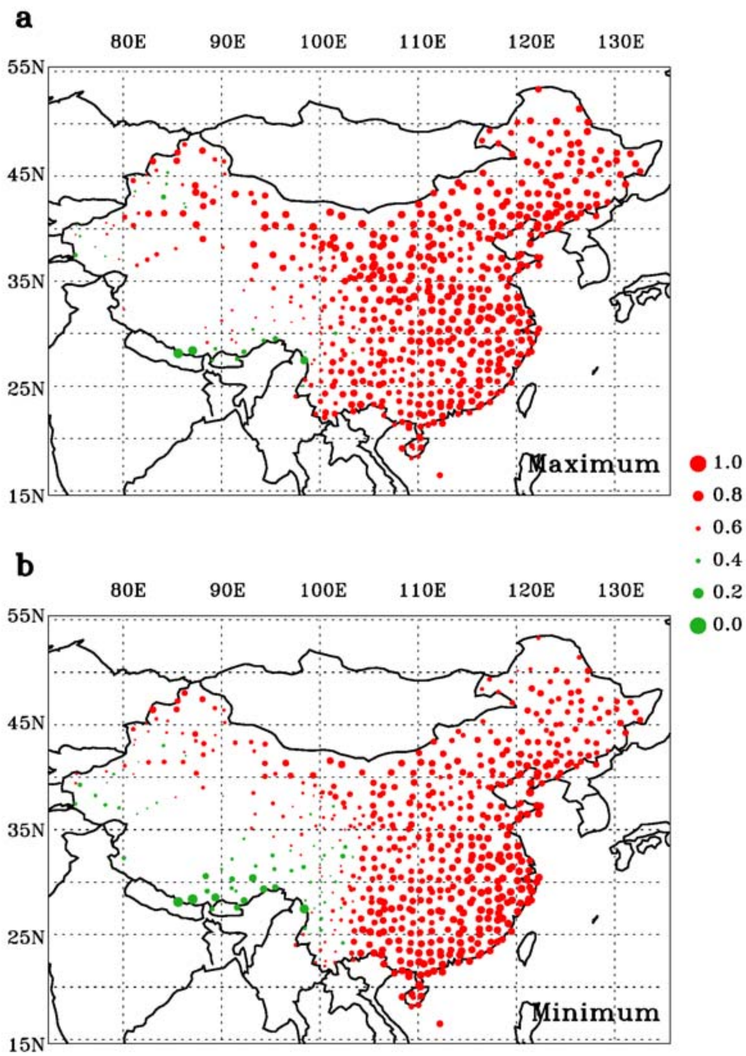

Fig. S5. Correlation coefficients between observed and R-2 monthly temperature anomalies for (a) maximum and (b) minimum during the period 1979-1998. The total sample is 240 for each station.

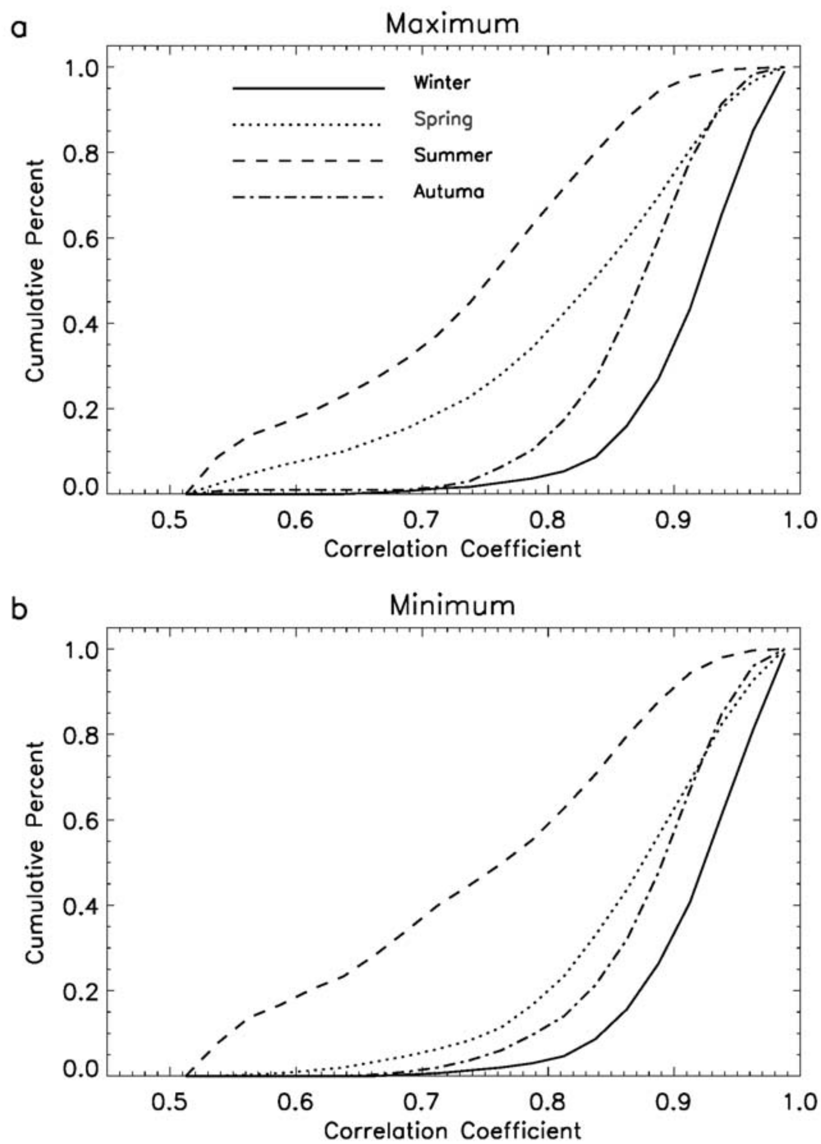

Fig. S6. Histograms of correlation coefficients between observed and R-2 temperature anomalies for (a) maximum and (b) minimum in winter, spring, summer, and autumn over our study region during the period 1979-1998. The total sample is 20 for each station.

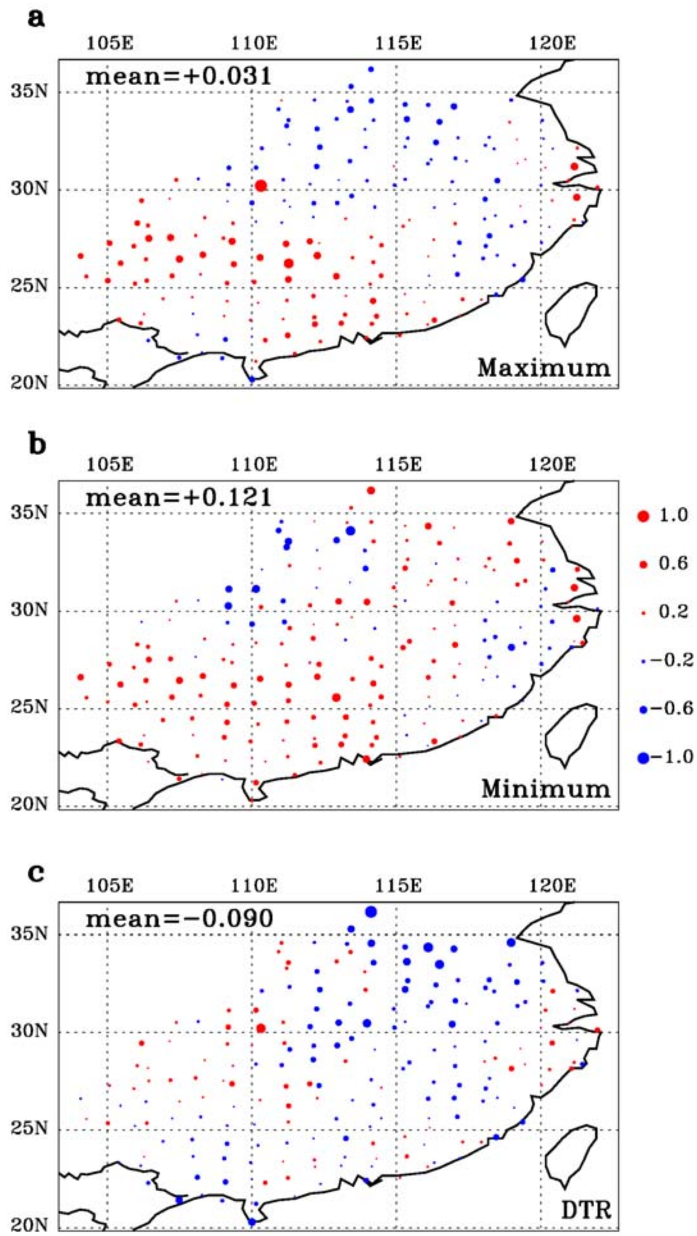

Fig. S7. Observed minus R-2 spring temperature trends (°C per decade) in southeast China during the period of 1979-1998: maximum (a), minimum (b), and DTR (c).

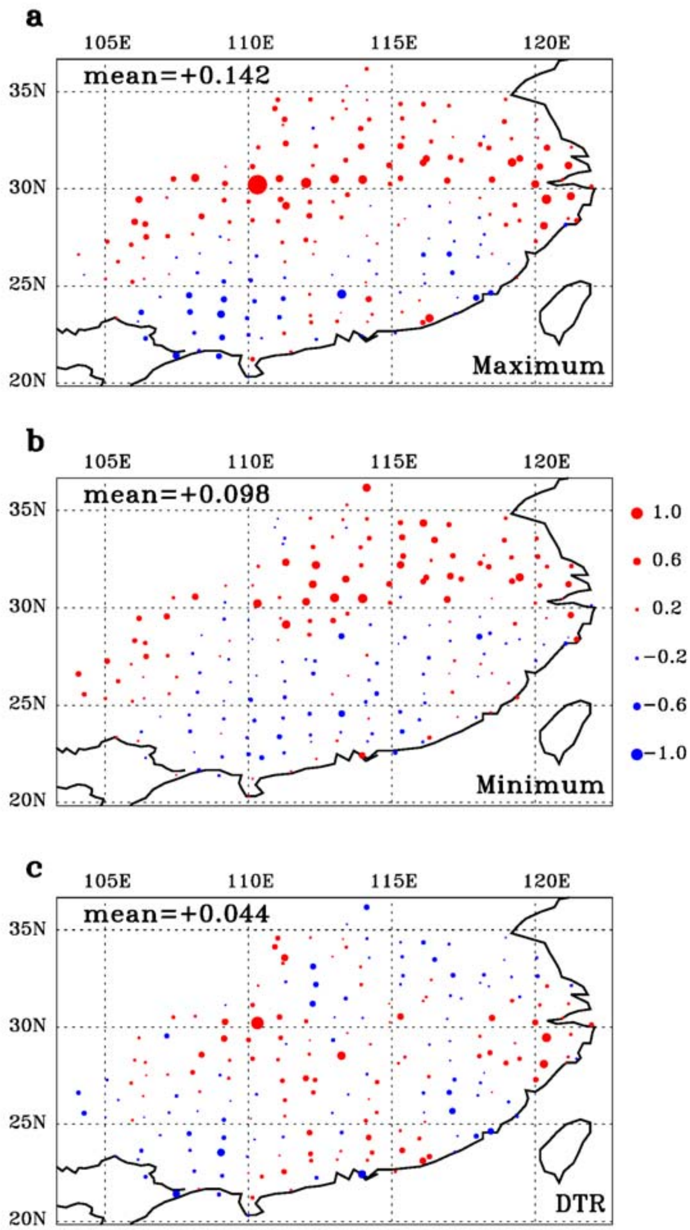

Fig. S8. Observed minus R-2 summer temperature trends (°C per decade) in southeast China during the period of 1979-1998: maximum (a), minimum (b), and DTR (c).

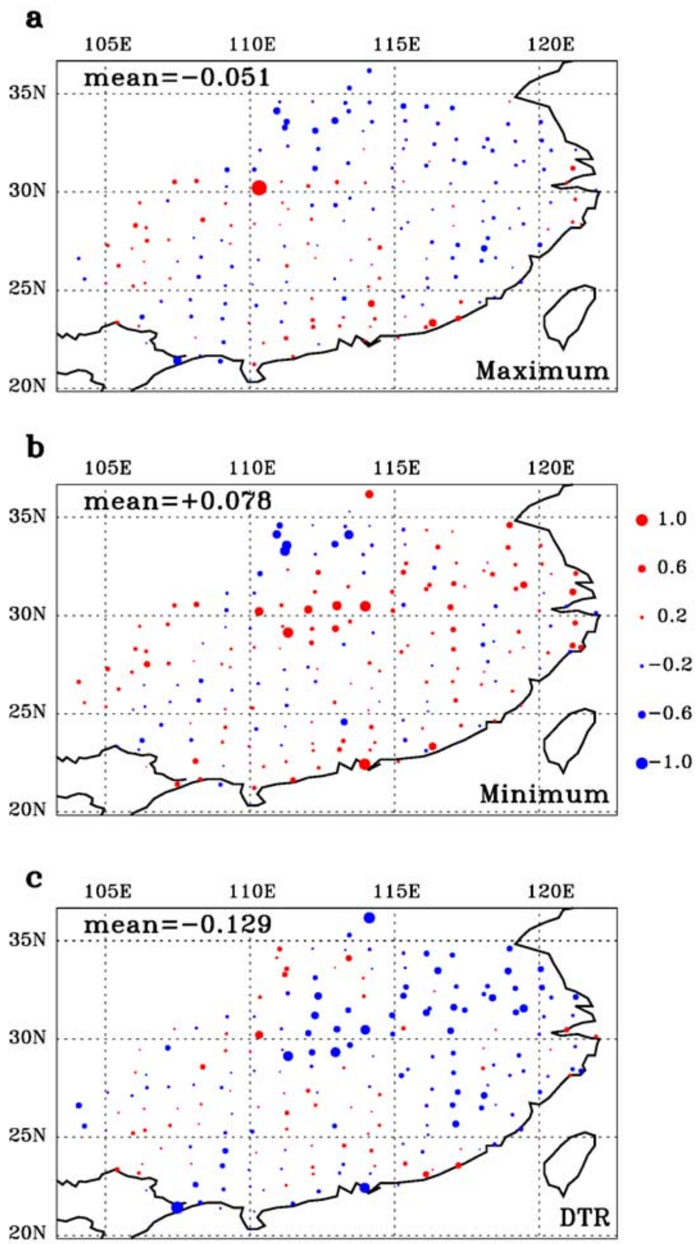

Fig. S9. Observed minus R-2 autumn temperature trends (°C per decade) in southeast China during the period of 1979-1998: maximum (a), minimum (b), and DTR (c).

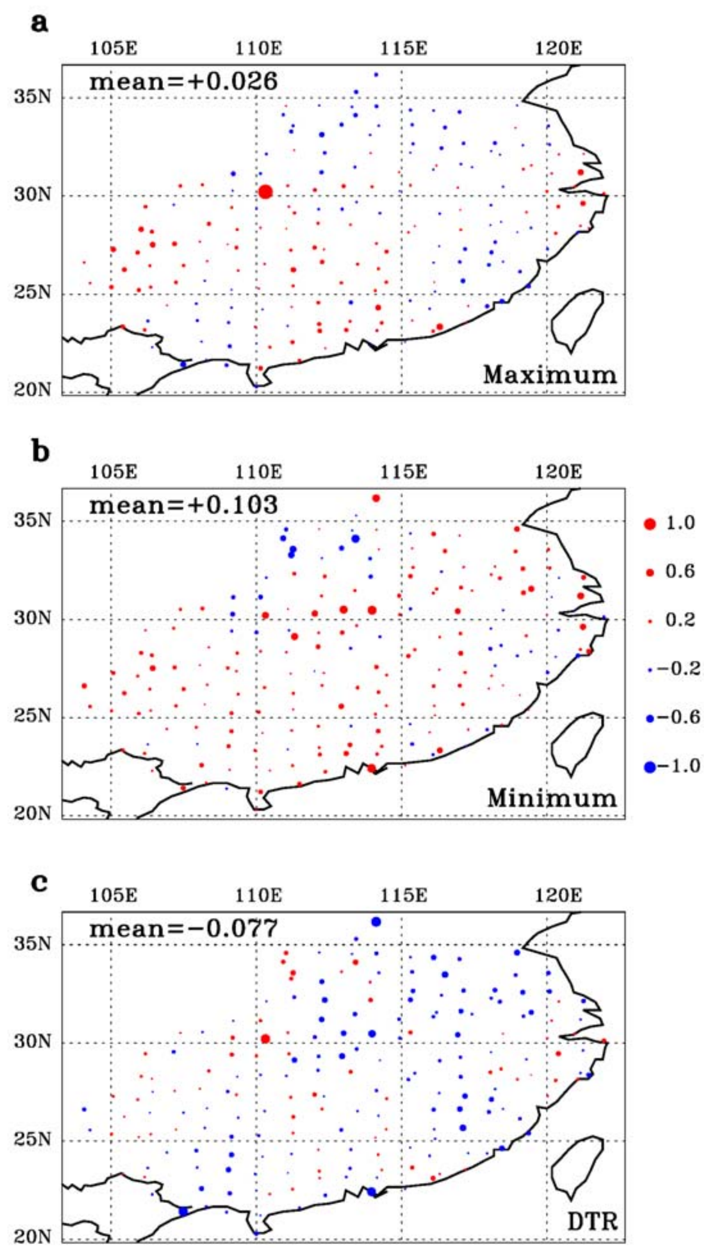

Fig. S10. Observed minus R-2 annual mean temperature trends (°C per decade) in southeast China during the period of 1979-1998: maximum (a), minimum (b), and DTR (c).