INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **24**: 853–870 (2004)
Published online 14 May 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/joc.1047

QUALITY CONTROL OF DAILY METEOROLOGICAL DATA IN CHINA, 1951–2000: A NEW DATASET

SONG FENG,[a] QI HU[a,*] and WEIHONG QIAN[b]

[a] *Climate and Bio-Atmospheric Sciences Group, School of Natural Resource Sciences, University of Nebraska-Lincoln, Lincoln, NE 68583-0728, USA*
[b] *Climate Dynamics Research Group, School of Physics, Peking University, Beijing 100871, People's Republic of China*

Received 3 December 2003
Revised 15 March 2004
Accepted 15 March 2004

ABSTRACT

Long-term observational data are essential for understanding local and regional climate and climate change. These data are also important for hydrological designs and agricultural decision making. This study examined the daily meteorological data from 726 stations in China from 1951 to 2000, and developed an unprecedented climatic dataset that contains 10 daily variables: maximum and minimum surface air temperatures, mean surface air temperature, skin surface temperature, surface air relative humidity, wind speed, wind gust, sunshine duration hours, precipitation, and pan evaporation. The characteristics of the original stations' data and quality-control methods designed and used in developing this dataset are detailed. The quality-control procedures identified less than 0.05% of the data records as being erroneous because of typos and incorrect units, reading, or data coding. When the spatial and temporal consistency of the variables' time series were inspected, nearly 37.9% of the stations were found to have one or more variables with inconsistent changes. The sources causing the temporal inconsistency/discontinuity were evaluated, and a method was developed and applied to adjust those data segments containing inconsistent values. The resulting data series, as an alternative to the original quality-controlled series, showed both spatially and temporally consistent trends in the occurrence frequency of extreme climate events compared with the unadjusted data series. Finally, the quality-controlled daily data were gridded to a $1.0^{\circ} \times 1.0^{\circ}$ grid system covering China after the erroneous and missing data were estimated. This new dataset opens up opportunities for analysing and understanding the climate variability and climate change in China. Copyright © 2004 Royal Meteorological Society.

KEY WORDS: China; new dataset; daily meteorological data; quality control

1. INTRODUCTION

Climate data are essential to our effort to identify and understand variations and changes of regional and global climate. Near-surface observations and radiosonde data also both provide the 'ground truth' for evaluating various remote sensing techniques and sensors, and for validating numerical model simulations. In recent years, the integration of the observation data and modelled data via assimilation techniques has yielded a new generation of datasets for advanced research purposes, e.g. the National Centers for Environmental Prediction–National Center for Atmospheric Research reanalysis data (Kalnay *et al*., 1996). These needs for climate data, in addition to applications of the data in local agriculture and water resources planning, have been propelling the development of quality datasets from ground and radiosonde observations, despite the recent multiple downsizes of many observation networks, such as those in the USA and Canada. As a result, daily and hourly datasets of various meteorological variables have been developed for many countries and regions in the world (Peterson *et al*., 1997; Manton *et al*., 2001; Klein Tank *et al*., 2002; Vincent *et al*., 2002).

* Correspondence to: Qi Hu, 237 L.W. Chase Hall, School of Natural Resource Sciences, University of Nebraska-Lincoln, Lincoln, NE 68583-0728, USA; e-mail: qhu2@unl.edu

Recently, a global daily meteorological dataset has been developed at the US National Climatic Data Center (NCDC; Gleason, 2002). In this global daily meteorological dataset, there are surface air temperature and precipitation data from 1951 to the 1990s at 196 stations in China (Gleason, 2002). This number is a small portion of the total number of stations in China (726) that have observation records from 1951 to 2000. In addition, most of those stations have measurements of other important meteorological variables, such as daily mean skin surface temperature (Ts; i.e. the temperature measured by a thermometer lying on bare ground, on top of a snow surface, or on a vegetation surface if it covers the ground, which is rather unique in the Chinese observing system), relative humidity (RH), wind speed W_s, wing gust W_g, sunshine duration hours (SS), and pan evaporation (PE). It is desirable to include the temperature and precipitation observations from the additional stations in the global dataset and, moreover, to establish a complete climate dataset not only of the conventional temperature and precipitation, but also of the other meteorological parameters for China for the period 1951–2000. To achieve the latter is the goal of this study.

This article summarizes the quality-control methods designed and applied to various data at 726 stations in China (Figure 1(a)) and describes a dataset of an array of meteorological parameters that will provide an

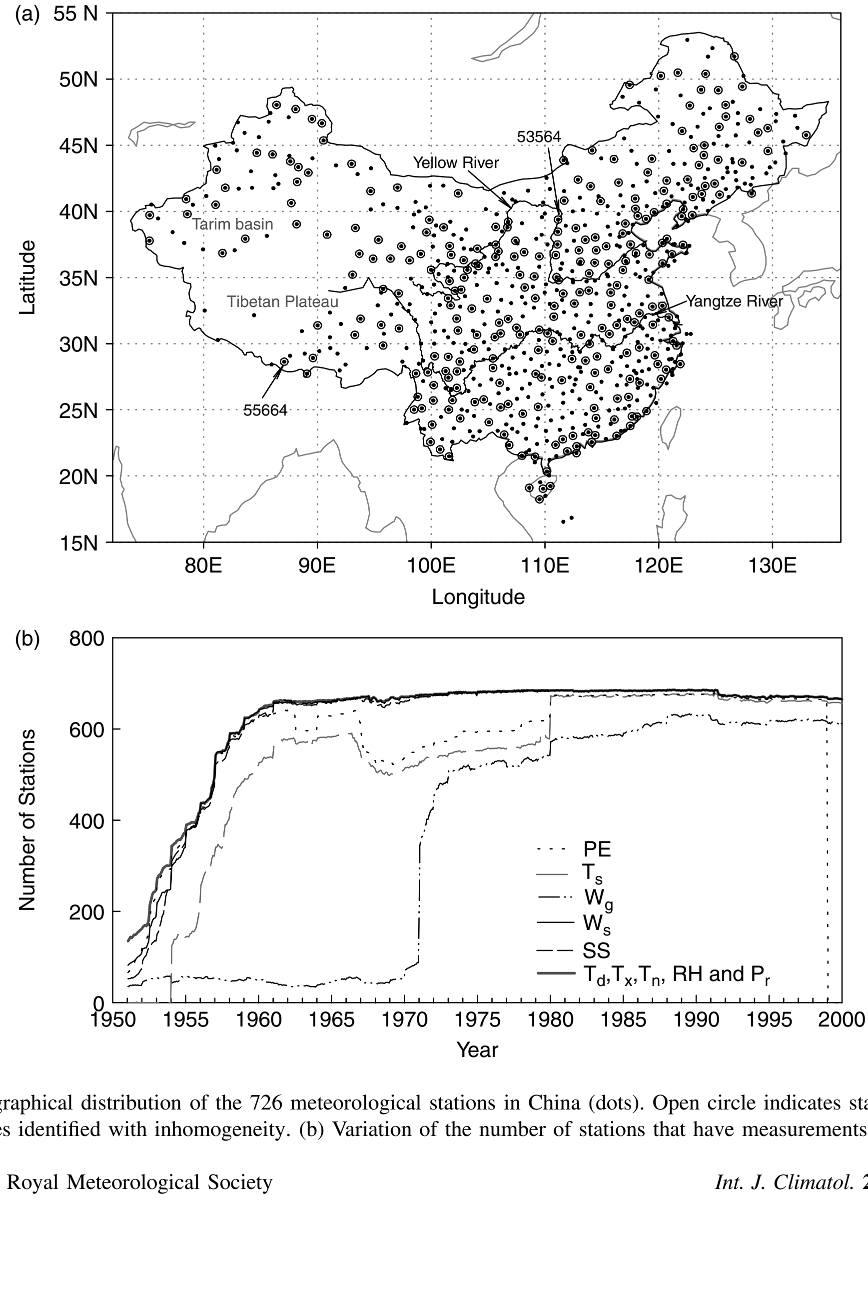

Figure 1. (a) Geographical distribution of the 726 meteorological stations in China (dots). Open circle indicates stations that have one of more variables identified with inhomogeneity. (b) Variation of the number of stations that have measurements for each variable

unprecedented resource for analysis and research of climate and climate change in China. The data sources and station characteristics are outlined in Section 2. Section 3 describes the quality-control methods. A method for checking consistency of daily data and methods to correct data errors caused by station moves and sensor upgrades are described in Section 4. In Section 5, we describe a gridded dataset derived after interpolation of the quality-controlled stations' data to a $1.0^{\circ} \times 1.0^{\circ}$ mesh covering China. Section 6 contains a summary and discussions of the uniqueness and advantages of the new datasets.

2. DATA

Daily data are obtained from the Chinese National Meteorological Centre (CNMC). These data are observations of 10 variables: daily maximum surface air temperature T_x, daily minimum surface air temperature T_n, daily mean surface air temperature T_d, daily total precipitation P_r, and the variables previously mentioned, i.e. T_s, RH, W_s, W_g, SS, and PE. Except for PE, which was measured in the period from 1 January 1951 to 31 December 1998, all variables are measured from 1 January 1951 to 31 December 2000. The standards of measurement instruments, except for PE sensors, have been outlined in Tao *et al.* (1991) and Kaiser *et al.* (1993), and are not repeated here. For PE, two kinds of evaporimeter were used in China. The PE presented in this study was measured by an evaporimeter with dimensions of 20 cm diameter and 10 cm height. The evaporimeter was placed on grass, measuring evaporation in a vegetation environment, similar to the Class-A PE measurement in the USA (Allen *et al.*, 1998). (Because of the difference in dimensions of the evaporation pans, measurements from the evaporimeter are expected to be different from the Class-A PE measurements. Their differences have never been carefully examined, however.)

The geographical distribution of the 726 stations is shown in Figure 1(a). The stations are fairly evenly distributed in the plains east of the 95 °E longitude. The distance between the stations increases and the coverage becomes coarser in the western and northern parts of China. A large void exists in the western Tibetan Plateau and the Tarim basin in Xinjiang province (the largest desert region in China). Although the number of stations in service has changed over the years, the total number of stations that measure T_d, T_x, T_n, T_s, SS, W_s and P_r has remained at about 726 after the network was established around 1958 (Figure 1(b)). The stations that measured the additional variables do vary, however. As shown in Figure 1(b), the number of stations measuring Ts, PE, and Wg has changed considerably, but has remained stable since 1980. There were very few stations having wind gust measurements prior to 1971. A change in 1971 has since upgraded the wind gust measurements.

3. QUALITY CONTROL

In this section, we describe the quality control and assurance that are designed and applied to identify erroneous data resulting from sensors and observation sources.

3.1. High–low extreme check for daily values

This method compares daily values of various variables from individual stations with established extreme values. Specifically, we use the temperature extremes defined by Kubecka (2001) and Gleason (2002) to check stations' daily temperatures, and we use the precipitation extremes given in Gleason (2002) to compare the stations' daily precipitation amounts. The wind-speed extremes are from Meek and Hatfield (1994), and the extremes for RH and PE are from the CNMC. For sunshine duration hours, we use zero for the low extreme and calculate the high extremes using the following equation (Allen *et al.* 1994):

$$H = \frac{24}{\pi} \times \arccos(-\tan\phi\tan\delta) \quad 0 \le H \le 24 \tag{1}$$

where $\delta = 0.409\sin(0.0172J - 1.39)$ is the solar inclination angle, in which J is the day of the year, with $J = 1$ being 1 January. In Equation (1), H is the maximum day length (hours) on day J, ϕ is the latitude of the station. Note that the calculated maximum day length is slightly different depending on the methods used (see Forsythe *et al*. (1995)).

These extreme values for each variable are summarized in Table I. Data values greater than (less than) the highest (lowest) values listed in Table I are flagged and are not used in subsequent quality-control calculations. The total number of errors from this check for the various variables is given in Table II. Among the list, T_s and SS have large numbers of flagged data from this check. For T_s, one source for such a large number of erroneous data has been incorrect coding of the data. This was particularly serious at several stations in northeastern China and the Tibetan Plateau during the winter season. A large number of erroneous data in SS also arose at a few stations.

3.2. Internal consistency check

Reek *et al*. (1992) outlined eight rules to identify erroneous data of air temperature and precipitation. They concluded that the errors are due to data reporting and digitizing, typos, unit differences, and the use of different based values in data reporting. In this study, we used three of their rules to check the daily data: (1) internal inconsistency, which identifies errors such as daily T_x being cooler than T_n, that the daily T_d is larger than T_x or less than T_n, and that daily W_s is larger than W_g; (2) excess diurnal temperature range check, which identifies errors with extraordinarily large daily temperature range ($T_x - T_n$) while T_x and T_n are within their reasonable ranges; and (3) a 'flat line' check, which identifies data of the same value for at least seven consecutive days (not applied to zero precipitation data). All the identified erroneous data are

Table I. High/low values used as extreme values of different variables

Variable	High	Low	Source
Mean surface air temperature T_d	93.9 °C	−89.4 °C	Kubecka (2001), Gleason (2002)
Maximum air temperature T_x	93.9 °C	−89.4 °C	Kubecka (2001), Gleason (2002)
Minimum air temperature T_n	93.9 °C	−89.4 °C	Kubecka (2001), Gleason (2002)
Mean skin surface temperature T_s	93.9 °C	−89.4 °C	Kubecka (2001), Gleason (2002)
Precipitation P_r	1828.8 mm	0 mm	Gleason (2002)
Sunshine duration (SS)	H	0h	Allen *et al*. (1994)
Wind speed W_s	45 m/s	0 m/s	Meek and Hatfield (1994)
Wind gust W_g	100 m/s	0 m/s	Possible extremes
Relative humidity (RH)	100%	0%	Physical constraints
Pan evaporation (PE)	500 mm	0 mm	CNMC

Table II. Percentages of erroneous and suspected data by categories. The total values are the sum of high/low extremes, internal consistency, spatial outlier, and temporal outliers that are larger than 5 BSD

	High/low extremes (%)	Internal consistency (%)	Temporal outliers (%)			Spatial outliers (%)	Total (%)
			>5 BSD	>4 BSD	>3 BSD		
T_d	0.0	0.00 048	0.010	0.042	0.375	0.0028	0.0133
T_x	0.0 00 035	0.00 073	0.016	0.052	0.393	0.0068	0.0236
T_n	0.0 00 009	0.00 021	0.012	0.046	0.365	0.0084	0.0206
T_s	0.0031	0.00 253	0.019	0.056	0.038	0.0064	0.0310
P_r	0.0	0.00 594	—	—	—	0.0213	0.0272
SS	0.0094	0.90 931	—	—	—	0.0319	0.9506
W_s	0.0	0.00 339	—	—	—	0.0764	0.0798
W_g	0.00 036	0.00 177	—	—	—	0.0133	0.0155
RH	0.0	0.00 390	—	—	—	0.0168	0.0207
PE	0.0	0.01 879	—	—	—	0.0730	0.0918

flagged. For erroneous data detected by the 'flat line' check, all the consecutive data are flagged as suspect values, except the first value. These rules are applied to individual station daily observations and the total numbers of suspected values are shown in the third column of Table II. It is evident from Table II that the SS and PE have many more errors identified from these checks than the other variables. (The large number of erroneous data of SS was suggested to be attributed to the fact that the SS measurement is a rather subjective matter of interpreting the burnt traces on the measurement cards, especially when the sun is out briefly and intermittently. Each observer makes somewhat different assessments. Additional errors also arise if the glass sphere of the recording device is not aligned precisely with the cards, or if the instrument is not well exposed and lacks a clear view of the sun.)

3.3. Temporal outliers check

The above checks identified some obvious erroneous values in the data series, but they cannot detect outliers that have the following problems: (1) where data values are much larger (or smaller) than neighbouring values but are not larger than the threshold for being detected by the consistency check; (2) where data create a large step change from the previous daily value(s). To identify these outliers, we use Lanzante's (1996) biweight mean and biweight standard deviation method. The biweight estimate is a weighted average such that weighting decreases away from the centre of the distribution. This method has also been used in Gleason (2002). According to Gleason (2002), we use the data from the day before and after the day having the suspect value of the current year, and the day before, after, and on the day from all other years in the same station to create a new time series X_i. For example, X for 6 June 1988 for a station with 50 years of data (1951–2000) would consist of the following daily observations: 5–7 June 1951–87; 5 and 7 June 1988; and 5–7 June 1989–2000. After we obtain the X_i series, the median M and absolute deviation from the median (MAD) are estimated. (The MAD is the median of the absolute deviations of the values from the median.) From the MAD, the weights u_i are calculated using

$$u_i = \frac{X_i - M}{c \times \mathrm{MAD}} \tag{2}$$

where c is a constant; c represents a 'censor' value, such that all observations beyond a certain critical distance from c are given zero weight. The censor value of 7.5 used by Lanzante (1996) is also applied in this study. In addition, we set $u_i = 1.0$ for any $|u_i| > 1.0$ to accomplish the censoring (Lanzante, 1996). With u_i, the biweight mean is estimated using

$$\overline{X_{\mathrm{bi}}} = M + \frac{\sum_{i=1}^{n}(X_i - M)(1 - u_i^2)^2}{\sum_{i=1}^{n}(1 - u_i^2)^2} \tag{3}$$

and the biweight estimate of the standard deviation is

$$s_{\mathrm{bi}} = \frac{\left[n\sum_{i=1}^{n}(X_i - M)^2(1 - u_i^2)^4\right]^{0.5}}{\left|\sum_{i=1}^{n}(1 - u_i^2)(1 - 5u_i^2)\right|} \tag{4}$$

Both $\overline{X_{\mathrm{bi}}}$ and s_{bi} are more heavily weighted towards the centre of their distributions than the wings or tails. Thus, they are more resistant to outlier values and provide a more robust estimation of mean and standard deviation than traditional statistical methods, which apply equal weight throughout the distribution (Lanzante, 1996). The $\overline{X_{\mathrm{bi}}}$ and s_{bi} are used to determine the Z-score of a particular day's observation using

$$Z = \frac{|X_{\mathrm{o}} - \overline{X_{\mathrm{bi}}}|}{s_{\mathrm{bi}}} \tag{5}$$

where X_{o} is the observed value. The Z-score values greater than or equal to three, four, or five biweight standard deviations (BSDs) from the biweight mean are classified with different flags to indicate different degrees of suspicion. According to Lanzante (1996), unless the sample size of X_i is at least 11 (18), it is not possible to obtain $Z > 3(>4)$. Therefore, all stations for which this type of outlier check is performed required at least 10 years of data with adequate numbers of observations centred at the date of a suspected value. Figure 2 shows the outliers identified from this method for a station in the Tibetan Plateau. All the flagged values with $Z > 3$ BSDs are much smaller (or larger) than the values collectively showing the annual cycle in the figure. The results in Figure 2 and the numbers given in the sixth column in Table II also indicate that a larger number of values are flagged when we use the criterion $Z \geq 3$ BSDs. Because $Z > 5$ BSDs indicates a data value that is 10 to 20 °C warmer or colder than the collective annual cycle (Figure 2), we use the criterion $Z \geq 5$ BSDs for all the data. However, the results indicate that this method is only applicable to the temperature data (T_{d}, T_{x}, T_{n} and T_{s}). The percentage of outliers in the data identified using this criterion is given in the fourth column in Table II.

3.4. Spatial outliers check

This method detects the outliers by comparing the data of neighbouring stations. Correlation coefficients R are computed for each month between daily data at a station (candidate station) and the 10 nearest stations. The minimum criterion is that R be significant at the 95% confidence level. Stations with large positive R are used to create their linear regression for the same variable between neighbouring stations and the candidate station. The root-mean-square error (RMSE) of the regressions is also computed. If more than five neighbouring stations have significant correlation with the candidate station at a specific month, than the five neighbouring stations with the lowest RMSE are chosen. After having N ($N \leq 5$) regression equations, we assign a daily value V_i of a variable to be suspicious if it falls outside the specified confidence intervals for all N pairs of stations (Hubbard, 2001):

$$\mathrm{VF}_{ij} - F \times \mathrm{RMSE}_j < V_i < \mathrm{VF}_{ij} + F \times \mathrm{RMSE}_j \tag{6}$$

where $j = 1, \ldots, N$ is the number of neighbouring stations, $i = 1, \ldots, m$, is the specific day in a month and m is the total number of days of that month, V_i is the data at the candidate station for day i, VF_{ij} is the fitted

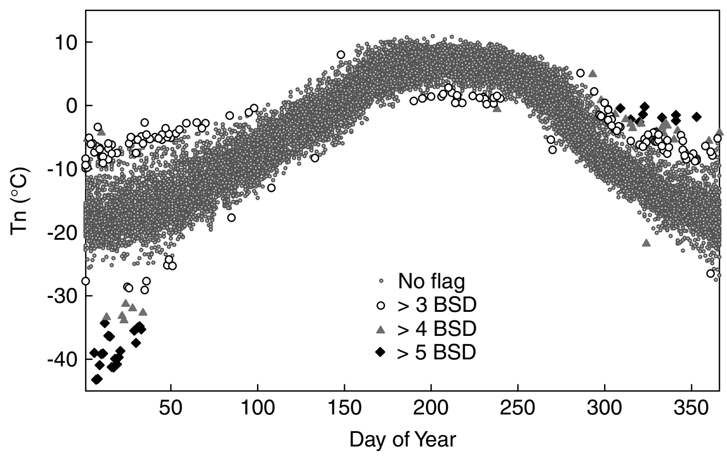

Figure 2. Temporal outliers of T_{n} identified for station 55 664 (28.63°N, 87.08°E) in the Tibetan Plateau

value by linear regression of neighbouring station j for day i, and F defines the desired confidence limit. In the present study, $F = 5$ for P_r and W_g and $F = 3$ for the other variables. The percentage of data failing this check for each variable is listed in the seventh column in Table II, and shows that less than 0.077% of the data were flagged from this check.

The total percentage of the suspicious or erroneous data identified from these previous four checks are shown in the last column of Table II. Except for SS, the daily data in China have high quality for most of the variables, which is comparable to that for the USA (Reek *et al.*, 1992).

3.5. Missing data

In our calculations of monthly values, the missing data (or data gaps) can induce temporal and spatial errors (Stooksbury *et al.*, 1999). Efforts have been made by some researchers to fill the missing data and correct the erroneous data before calculating the monthly mean (Eischeid *et al.*, 2000; Hubbard, 2001). In the present study, the suspicious data screened by the four previous checks and missing data are estimated using the following method (Hubbard, 2001)

$$v_{ei} = \sum_{j=1}^{N} \left[\mathrm{VF}_{ij} \times \mathrm{RMSE}_j^{-2} \right] \Big/ \sum_{j=1}^{N} \mathrm{RMSE}_j^{-2} \tag{7}$$

where v_{ei} is the estimated value and the other symbols are the same as in Equation (6).

We should point out that the number of neighbouring stations used in Equations (6) and (7) is not fixed in time. The number varies depending on the availability of station data for the year/month in question. Accordingly, the regression models also change in time. Moreover, the surrounding stations that may be optimal for a particular calendar month (e.g. January) may not be optimal for a different month (e.g. July). Thus, the spatial outlier check and estimation of missing data are constructed and applied for individual calendar months.

4. HOMOGENEITY CHECK AND HOMOGENIZATION

Yan *et al.* (2001) used individual stations' metadata and showed sudden changes of the mean temperature at two weather stations in China after their geographical locations changed. Similar changes are expected in this dataset for stations that have experienced location changes and/or sensor upgrades. These changes could have created erroneous trends or amplitudes of daily or monthly variations of temperature, for example, and must be identified and their corrections estimated. An ideal way to accomplish this task is to use the individual station's metadata that record the move and sensor changes. After identifying those changes we can compare the temporal variations of the data at the same station and the variations with that at neighbouring stations and examine the effect of a location move on the station's data homogeneity ('homogeneity' means that variations in the data have a consistent 'tempo'). However, there is very little metadata available for the stations in China: only 60 out of the 726 stations (Tao *et al.*, 1991; Kaiser *et al.*, 1993). Thus, alternative subjective and objective methods are needed. Numerous methods have been used to evaluate the homogeneity of monthly and annual climate data series (Jones *et al.*, 1986; Folland and Parker, 1995). A comprehensive review on those methods is given in Peterson *et al.* (1998). In China, Wang and Gaffen (2001) used subjective visual inspections to identify data inhomogeneities. Feng (1999) used three statistical methods, i.e. moving t-test (Peterson *et al.*, 1998), standard normal homogeneity test (Alexandersson, 1986), and departure accumulating method (Buishand, 1982) to check the homogeneity of the data series. His results compared favourably at stations with metadata information in western China. Because such a combination of statistical methods is often considered most effective to uncover data inhomogeneity (Wijngaard *et al.*, 2003), we use these three objective statistical methods to check the homogeneity in this work.

4.1. Create the reference series

A change in a station's time series may indicate inhomogeneity, but it may also simply be a result of an abrupt change in the region's climate (Peterson *et al.*, 1998). To isolate the effect of station change from regional climate change on data homogeneity, several techniques have been introduced. Most of these techniques use data from nearby stations to establish a descriptor of the regional climate (reference time series). In this study, the technique proposed by Peterson and Easterling (1994) is used to create the annual reference time series. In doing so, the difference series of a variable, e.g. for $V, \mathrm{FD}_i = V_{i+1} - V_i$ for the ith year is created and the mean of the central three FD values of the five closest neighbouring stations is used to create the difference reference series. (This procedure is designed under the belief that if there is a significant discontinuity in one of the five stations in a year, then that station would most likely have the highest or lowest FD value.) The difference series is then transformed into a new time series using $V_1 = 0$; $V_{i+1} = V_i + \mathrm{FD}_i$. This new time series is further adjusted so that the final year's value of the series is equal to the final year's value of the candidate series. After this adjustment, we obtain the station's reference time series. This method has been used for both temperature (Peterson and Easterling, 1994) and PE time series (Lawrimore and Peterson, 2000), and is adapted in this study to create the reference series of all the variables. The annual average value of each variable, except for precipitation, is used directly to create its reference time series. For precipitation, we first normalized it and then create its normalized reference time series. The reference series of precipitation is obtained by taking the product of S_1 and S_2, where S_1 is the standard deviation of the annual total precipitation at the candidate station and S_2 is its normalized reference time series.

4.2. Homogeneity check

The reference time series are used in checking the data homogeneity for individual variables. In this procedure, a time series Y is computed for a variable with Y_i being the difference between the candidate station data value and its reference value. For precipitation, Y_i represents the ratio of the candidate station data and its reference value. The three statistical methods used in Feng (1999) are applied to test the Y series. If Y has a big jump in year p and rejects the null hypothesis of no jump at the 99% confidence level, then the target station is suspected to be inhomogeneous in year p. In this study, a station's data are considered suspect when all three of the statistical tests reject the null hypothesis at the 99% confidence level.

The results of the three tests applied to station 53 564 (39.38 °N, 111.15 °E, see Figure 1) are shown in Figure 3 with the annual values of T_d, T_x, T_n, RH, PE, W_s, and W_g and their reference and Y series. A sudden change in 1973–74 is clearly shown in all the variables except for Wg. The failure of the statistical tests to find the change in W_g is mainly because the change occurred near the beginning of the W_g time series (see the last two panels in Figure. 3). The station's metadata suggested that the station was moved from a location at 39.28 °N, 111.27 °E to a site at 39.38 °N, 111.15 °E on 1 January 1974, a record consistent with the abrupt change in the station's data. The relocation apparently had a stronger effect on T_n than on T_x, as depicted by the larger differences between T_n and its reference time series in Figure. 3. The effect is also weak on T_d, and nearly diminishes on T_s (Figure 4). Figure 4 also suggests a small effect of the relocation in 1974 on P_r and SS. This weak effect on SS may be attributed to the latitude of the station changing by only $10'$ in the relocation, because SS is a heavy function of a station's latitude; see Equation (1). The small effect on P_r could be partially attributed to the fact that P_r has large seasonal and interannual variations, which may have shadowed P_r changes caused by the relocation.

Because the three statistical methods used in the procedure have different sensitivities to different changes in a station's data series, the test results from these methods sometimes have discrepancies. For example, two methods may suggest an abrupt change at a time in the data series but the third may show little support for such a change. Similar problems were also reported with other methods (Wijngaard *et al.*, 2003). To resolve this uncertainty without a station's metadata, we visually evaluate the data series from those stations with 'disputed' results from the three statistical methods and subjectively decide whether the changes detected in the data series have occurred. This hybrid testing method allowed us to detect even small effects of a station's relocation and other changes on the data series homogeneity, albeit that not all effects from changes of similar

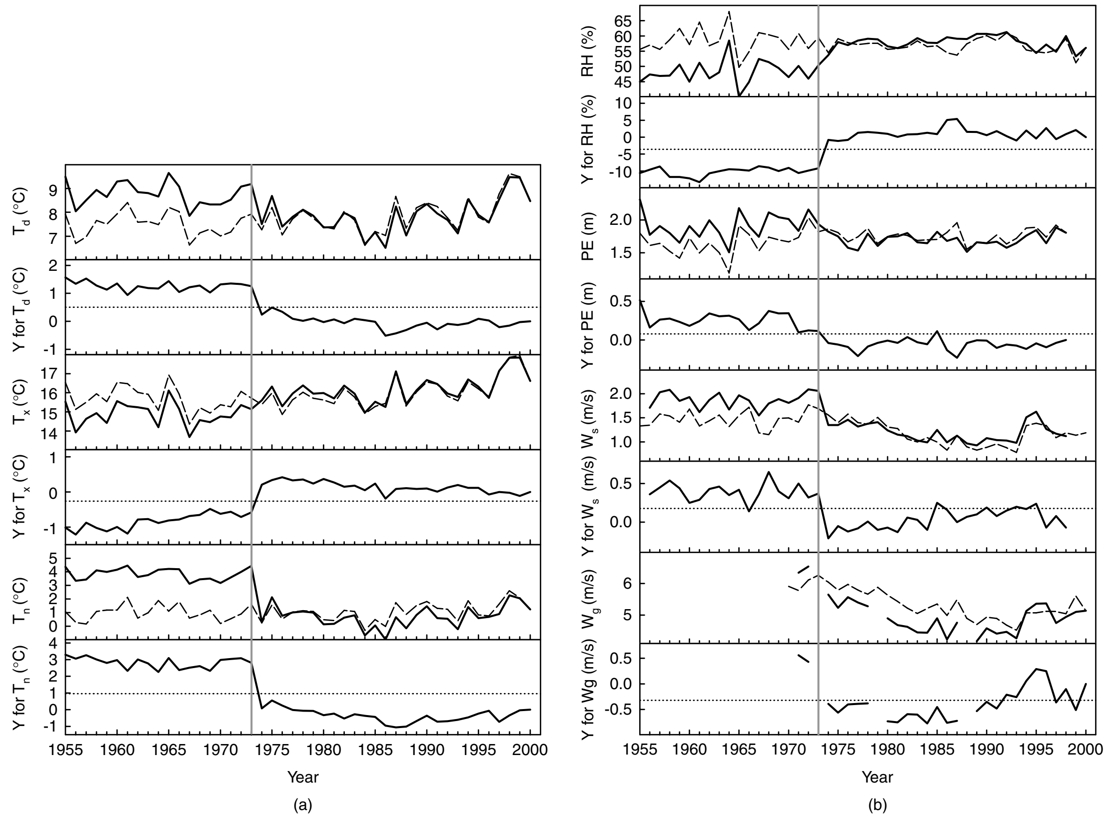

Figure 3. Variations of measured (thick line), reference (dash line), and the Y series of T_d, T_x, T_n, RH, PE, W_s, and W_g for station 53 564 (39.38°N, 111.15°E). An interruption in the homogeneity in 1973 is shown in each variable at the station

sorts at a station could be detected with certainty. The total number (also in percentage) of stations whose data were found having inhomogeneities in the variables is given in Table III. Up to 11.4% of stations have inhomogeneities in their T_d, T_x, T_n, W_s, and PE data. Overall, nearly 37.9% of the stations have one or more variables suffering abrupt changes.

In Table III, many stations that have inhomogeneities in other variables do not show a similar inhomogeneity in W_g. Besides the relatively short W_g observations compared with the other variables (Figures 3 and 4), this difference may suggest a large-scale feature of strong wind in those regions where the W_g showed no change from relocation of stations. Additionally, precipitation is also less sensitive to station relocation and sensor upgrade than temperature, a result consistent with the high variability in precipitation (Wijngaard *et al.*, 2003).

4.3. Homogenization

When a station's data are identified as inhomogeneous, that station's data series then becomes questionable or invalid for trend and variability analysis. Given that, overall, 37.9% of the station's data series have inhomogeneities, it is desirable to 'remove' the abrupt changes so that the data series could be useful, or have

Table III. Number and percentage of stations with identified inhomogeneity for each variable

	T_d	T_x	T_n	T_s	P_r	RH	SS	W_s	W_g	PE
Number	77	66	82	30	12	57	21	83	7	65
Percentage	10.6	9.1	11.3	4.1	1.7	7.9	2.9	11.4	1	9.0

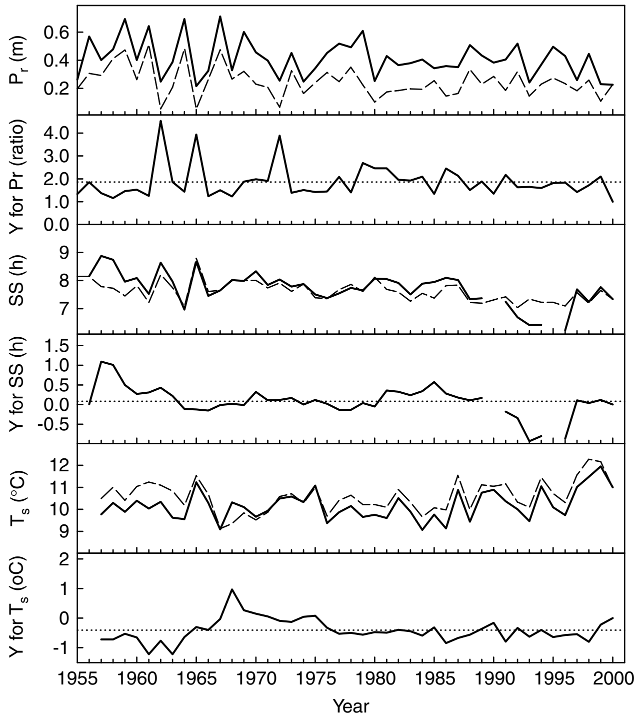

Figure 4. Variation of measured (thick line), reference (dash line), and the Y series of P_r, SS, and T_s for station 53 564

some reference value, to climate studies in China. We note that this effort is to 'rescue' some station's data for potential use in research and is not intended to create a new data series. In the dataset, these data are presented separately from the original data series, with flags showing the abrupt changes and inhomogeneity.

An ideal 'rescue' method is to assess each cause of the abrupt change in the data series and quantify its effect on daily data after the abrupt change (Yan *et al.*, 2001). Lacking adequate metadata often makes this method impractical, however. In this study, we use the method outlined in Vincent *et al.* (2002) to minimize the inhomogeneity in the data series (except for P_r and W_g). When a station's data series has inhomogeneity, the five nearest neighbouring stations with homogeneous data are selected to create a reference series for each month using the method outlined in Section 4.1. Then, Vincent *et al.*'s (2002) method is used to compute a monthly adjustment for a variable

$$\mathrm{VM}_i = a + bI + c\mathrm{VR}_i + e \quad i = 1, \ldots, N \tag{8}$$

where $I = 0$ for $i = 1, \ldots, p$, and $I = 1$ for $i = p + 1, \ldots$, N; p is the year when an abrupt change occurred and N is the total number of years in the data. In Equation (8), VM_i and VR_i are the monthly observations at the target station (which has the inhomogeneity) and the reference series respectively. The coefficients a, b, and c are determined statistically. Monthly corrections are obtained by applying Equation (8) to time series of the 12 individual months. The monthly corrections so obtained correspond to the magnitude given by the parameter b for each month. As an example, the monthly corrections for T_x, T_n, W_s, and RH of station 53 564 are shown in Figure 5. For these variables, the corrections vary from month to month. They also vary between variables, indicating different environment conditions at the locales of the station.

According to Vincent *et al.* (2002), the daily adjustments can be derived from the monthly corrections using a linear interpolation between midmonth 'target' values that are objectively chosen so that the average of the daily adjustments over a given month is equal to the monthly correction. The 'target' values are related to the monthly correction by the following matrix relationship (Vincent *et al.*, 2002):

$$\mathbf{T} = \mathbf{A}^{-1}\mathbf{M} \tag{9}$$

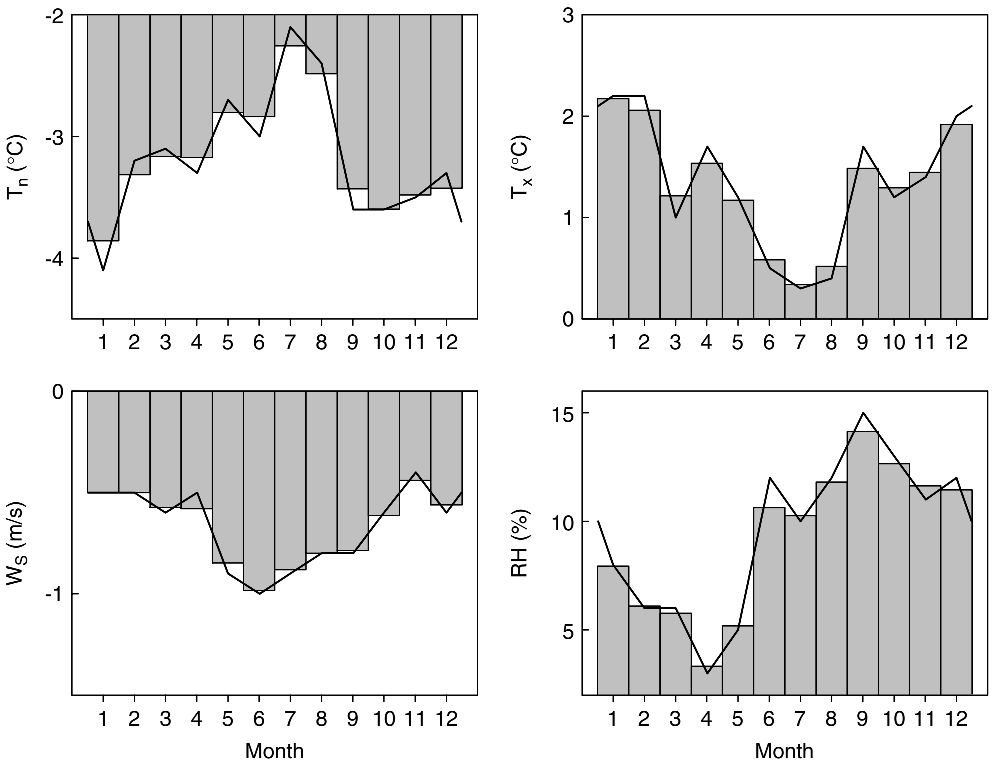

Figure 5. Monthly (bars) and daily (lines) adjustments for T_n, T_x, W_s, and RH for station 53 564

where $\mathbf{T}$ is a 12×1 vector consisting of the target values, $\mathbf{M}$ is a 12×1 vector consisting of monthly corrections, and $\mathbf{A}$ is a tridiagonal 12×12 matrix:

$$\mathbf{A} = \left| \begin{array}{ccccc} 7/8 & 1/8 & & & \\ 1/8 & 6/8 & 1/8 & & \\ & . & . & . & \\ & & 1/8 & 6/8 & 1/8 \\ & & & 1/8 & 7/8 \end{array} \right|$$

The target values obtained by Equation (9) are assigned to the middle day in each month and are then linearly interpolated to obtain the daily adjustments in the month. In this procedure, the average of the daily adjustments over a given month is equated to the monthly correction. As shown in Figure 5, the daily adjustments (the solid line) have captured well the monthly corrections for T_x and T_n. Figure 5 also shows the daily adjustments for W_s and RH. By capturing the monthly corrections, these results suggest that Vincent *et al.*'s (2002) daily adjustment method is not only accurate for daily temperature adjustment, but is also accurate for other variables such as SS, PE, W_s, and RH.

After calculating the daily adjustments we added them to the daily values before the year with the abrupt change, i.e. year p (correcting to the current station's environment). During the adjustment, we also make the following adjustments: when a new daily value of SS, PE, W_s, or RH is less than zero, it is set to zero; when RH is larger than 100% it is set to equal to 100%; excessive values of SS are set to the possible maximum duration hours.

4.4. Impact of the adjustments on calculated climate trend

While removing the abrupt changes in a data series, the adjustment procedures could also have introduced errors to the adjusted data series, because the adjustment procedures are not necessarily describing the actual changes. Potential errors induced by the daily adjustments have been discussed by Vincent *et al*. (2002). Their results from comparisons of unadjusted and adjusted daily temperatures from stations in Canada showed that the adjustments improve the data consistency: 'The adjusted daily temperatures were closer to the actual observations and exerted little effect on the temperature standard deviation'. The adjusted series also show frequency and distribution of temperature extremes closer to the observed than the unadjusted data series. In this section, we show the differences of trend in frequency and extreme weather events, e.g. severe cold days, windy days, and high-humidity days, between the estimated and the original data series. Our purpose is to illustrate the improvement by the homogeneity adjustment to the data at stations that experienced abrupt changes, and should not be understood as an encouragement for use of the adjusted data in any means. Such a decision is left for the users.

The seasonal trends in the occurrence frequency of severe cold days, windy days, and high-humidity days in 1958–2000 is shown in Figures 6–8 for the stations whose data have inhomogeneities and have been adjusted. The number of severe cold days (or cool days in summer) in each year is determined by the percentage of days with T_n colder than the 5th percentile of the period 1961–90, and the number of windy days and high-humidity days is determined by the percentage of days with W_s and RH respectively higher than the 95th percentile. Comparisons of each pair of the panels in Figures 6–8 indicate rather random trends in changes of those variables in the unadjusted data. Large trends of opposite changes are observed at stations next to each other, suggesting spatial inconsistency of the data at those stations. Such inconsistency is not seen in the distribution of the adjusted trend. The adjusted data show positive trends in numbers of cold days during winter and spring (Figure 6), a result consistent with large winter warming in China (Qian *et al*., 2001). In summer, a rising trend in the numbers of cool days appears in the east Tibetan Plateau and the middle and lower reaches of the Yangtze River valley. These changes are in agreement with a cooling trend in summer minimum temperatures (Shen and Varis, 2001). In Figure 7, the adjusted data for windy days show a trend of decreasing number of windy days in recent decades, except for the Tibetan Plateau. The decreased number of windy days in summer is consistent with the weakening seasonal mean southerly winds since the late 1970s found by Qian and Zhu (2001). Figure 8 shows a trend of an increasing number of summer high-humidity days in the Yangtze River valley and a decreasing trend in northern China, suggesting a wetter climate in the Yangtze River valley and drier conditions in northern China (Hu and Feng, 2001).

These improvements in describing the trends of various meteorological parameters and variables by the quality-controlled data (Figures 6–8) can be attributed largely to the fact that the quality controls have improved the continuity and consistency of the data series, after removing the erroneous data and replacing them with the estimates obtained under the constraint of maintaining the continuity in the natural variations of those parameters.

5. A GRIDDED DATASET

Because gridded daily datasets are easy to use for climate research and for applications in management and decision making, we develop a gridded daily dataset from the quality-controlled and adjusted stations' data for 1951–2000. Average data values of the variables in each 1.0° latitude × 1.0° longitude box are calculated using a modified Cressman (1959) scheme (Glahn *et al*., 1985; Charba *et al*., 1992). The same scheme was used in developing the gridded daily temperature and precipitation datasets in the USA (Higgins *et al*., 1996; Janowiak *et al*., 1999), and has been used for similar purposes at the US NOAA Climate Prediction Center. In applying this scheme, we first estimate daily values at the centre of each 0.5° × 0.5° grid-box and then take the area-weighted averages of the four 0.5° × 0.5° estimates to obtain the area-mean daily value for each 1.0° × 1.0° box. Because the spatial density of the stations is low in western China, particularly in the west Tibetan Plateau and the Tarim basin (Figure 1), a 1.0° × 1.0° latitude/longitude box is tenuous in those regions. As with any objective analysis procedure, our data are smoothed and the analysis bleeds into regions where no observations exist. Janowiak *et al*. (1999) demonstrated that the station reporting frequency difference and missing data might cause a jump (or drop) in gridded data values between grids, especially in mountainous regions. Therefore, use of the gridded daily data in the Tibetan Plateau should be undertaken with caution.

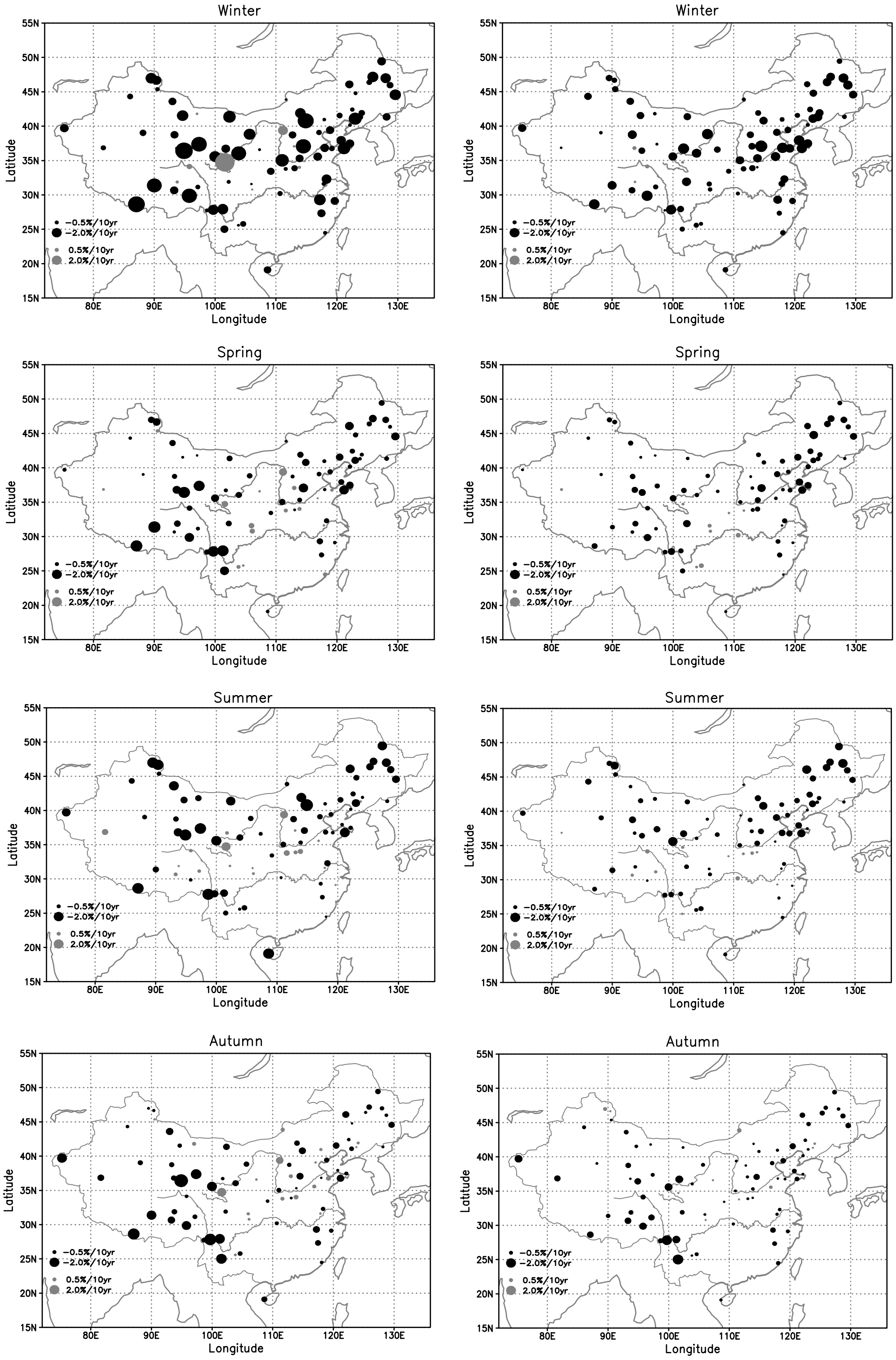

Figure 6. Trends in the percentage of days with minimum temperature colder than the 5th percentile value of each season over the period 1958–2000. The panels on the left show the results before the adjustment and the panels on the right are the counterparts after the adjustment

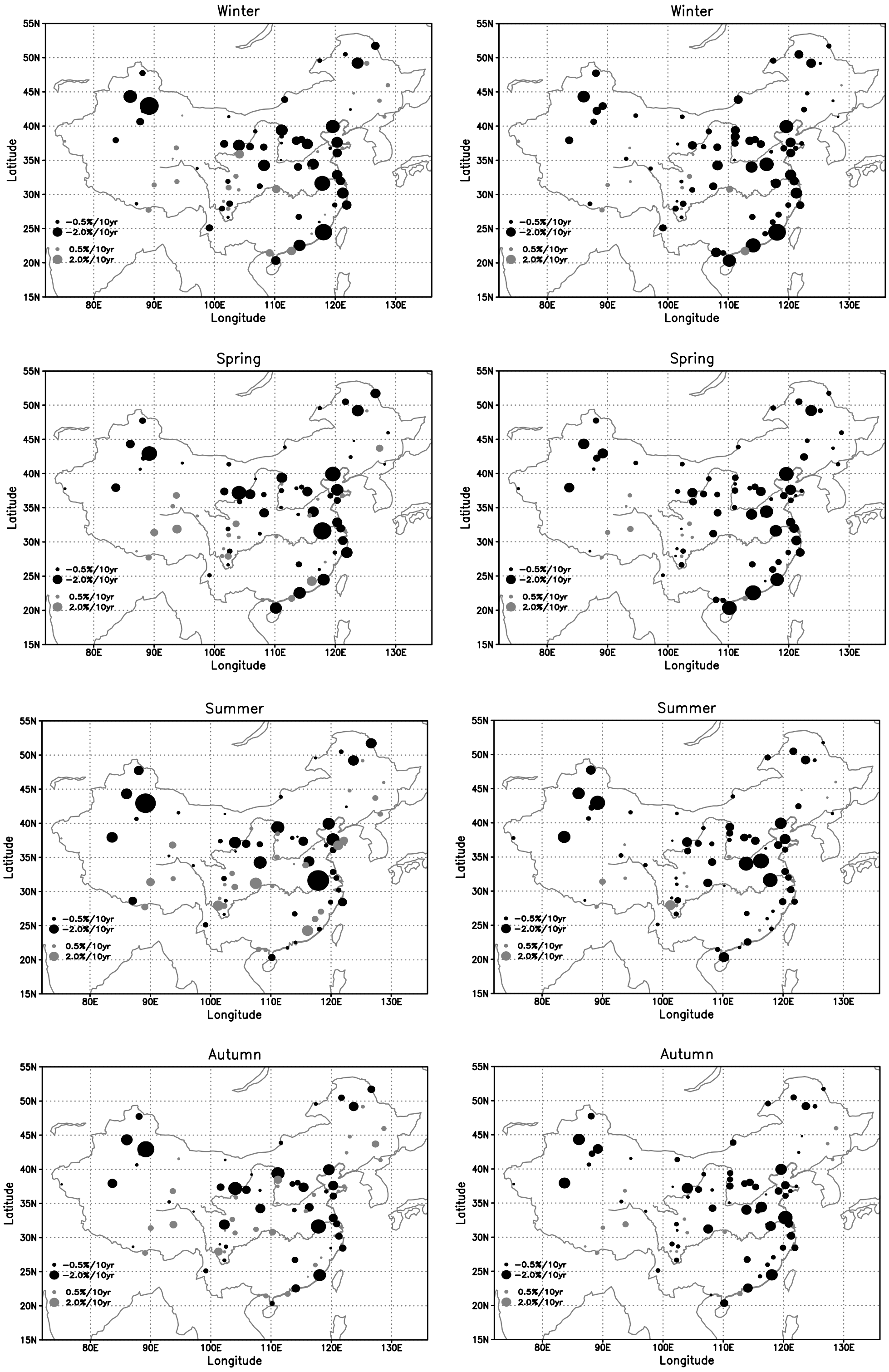

Figure 7. Trends in the percentage of days with daily mean wind speed stronger than the 95th percentile value of each season over the period 1958–2000. The panels on the left show the results before the adjustment and the panels on the right are the counterparts after the adjustment

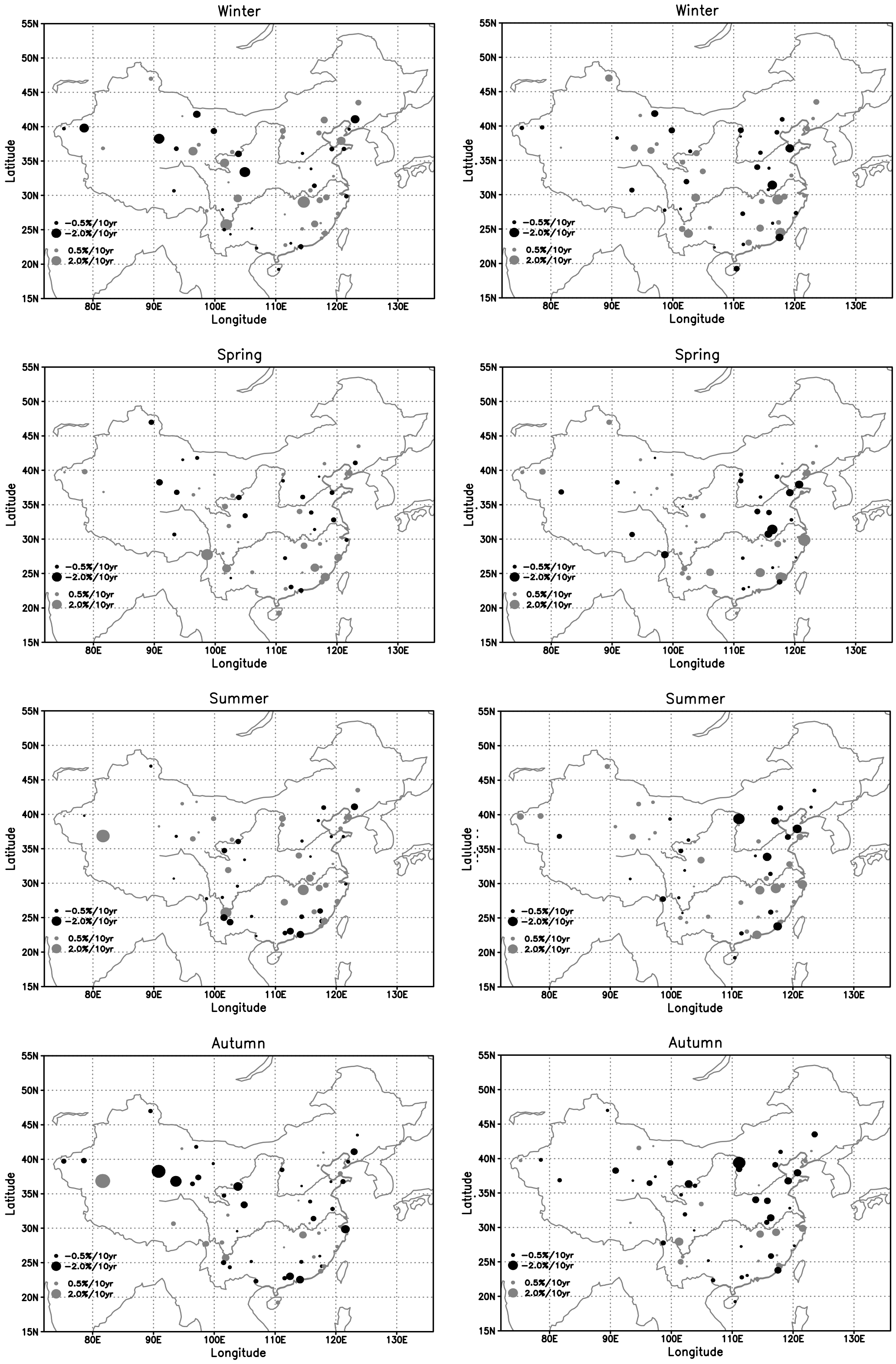

Figure 8. As for Figure 7, but for daily mean relative humidity

Nonetheless, as elaborated in Janowiak *et al*. (1999), such discontinuities should have been minimized in the gridded dataset because the Cressman (1959) scheme applies information from stations outside of a particular grid-box when estimating the grid value in that box.

6. SUMMARY

An unprecedented daily meteorological dataset of 1951–2000 is developed for China. The raw data used to build this dataset are individual stations' data from 726 stations in China. Each station's data contain the 10 variables T_d, T_x, T_n, T_s, P_r, SS, RH, PE, W_s, and W_g over the period from 1951 to 2000. These stations' daily data are examined for consistency in both temporal and spatial variations. Outliers are flagged. A total of less than 0.05% of the daily values is identified as erroneous and flagged. After the consistency check, three statistical methods are used to check the inhomogeneity in the data series that may be caused by relocation of stations and/or sensor problems. Data discontinuities resulting from such interruptions are identified by these methods, and are confirmed using the stations' metadata where available. Because of very limited metadata, a visual inspection is performed to determine data discontinuities when the three statistical methods yield different and ambiguous results for discontinuities in the data series.

A method developed by Vincent *et al*. (2002) is modified and used to eliminate the inhomogeneities and adjust the daily series. Magnitudes of the daily adjustments vary from month to month and from variable to variable. Comparisons between the unadjusted and adjusted daily series indicate that the general spatial pattern of trends in the occurrence frequency of extreme events and extreme values is similar, but the data after the adjustments display consistent details in spatial variations of trends and, thus, provide an improvement to the dataset. The differences between the adjusted and unadjusted data series also suggest that the Vincent *et al*. (2002) method can be applied not only to daily temperatures, but also to other daily meteorological variables. The adjusted stations' data are included in the dataset, as an alternative to the quality-controlled data, and a gridded dataset of $1.0° \times 1.0°$ resolution has been developed using the adjusted data. We understand that the adjustment may also induce errors in the data, and the decision on whether to use the adjusted data is left for the users.

One of the unique features of this dataset is that it has more than just the 'traditional' climate variables, i.e. the surface air temperature and precipitation. By including the daily variables of atmospheric humidity, wind, sunshine duration, and pan evaporation, this dataset opens opportunities for studying various aspects of climate variability and climate change in China.

ACKNOWLEDGEMENTS

We would like to thank Dr Lucie A. Vincent, Meteorological Service of Canada, for providing the code for daily data adjustments, and Dr Kenneth G. Hubbard of the University of Nebraska–Lincoln for useful discussions during this work. This work was initiated in a collaboration of Dr Q. Hu and Dr W. Qian, and completed while Q. Hu was a visiting scientist at the Department of Atmospheric Science of Peking University. Both the original and quality-controlled datasets reside at Peking University. Users of the quality-controlled dataset should contact W. Qian (qianwh@pku.edu.cn). During the collaboration, W. Qian was supported by China's National Key Program for Developing Basic Sciences (no. G1999043405), and Q. Hu was supported by the USDA Cooperative Research Project NEB-40-008.

This paper forms part of Agricultural Research Division, University of Nebraska-Lincoln Contribution Number 14581.

REFERENCES

Alexandersson H. 1986. A homogeneity test applied to precipitation data. *Journal of Climatology* **6**: 661–675.

Allen RG, Smith M, Peerier A, Pereira LS. 1994. An update for the calculation of reference evapotranspiration. *ICID Bulletin* **43**: 35–92.

Allen RG, Pereira LS, Raes D, Smith M. 1998. Crop evapotranspiration: guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper No. 56. FAO, Rome.

Buishand TA. 1982. Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology* **58**: 11–27.
Charba JP, Harrell III AW, Lackner III AC. 1992. A monthly precipitation amount climatology derived form published atlas maps: development of a digital database. TDL Office Note 92–7. NOAA, US Department of Commerce.
Cressman GP. 1959. An operational objective analysis system. *Monthly Weather Review* **87**: 367–374.
Eischeid JK, Pasteris PA, Diaz HF, Plantico MR, Lott NJ. 2000. Creating a serially complete, national daily timeseries of temperature and precipitation for the western United States. *Journal of Applied Meteorology* **39**: 1580–1591.
Feng S. 1999. Comprehensive analysis of Qinghai–Xizang (Tibet) Plateau climate fluctuations on decadal to millennial timescales. PhD dissertation, Lanzhou Institute of Plateau Atmospheric Physics, Chinese Academy of Sciences.
Folland CK, Parker DE. 1995. Correction of instrumental biases in historical sea surface temperature data. *Quarterly Journal of the Royal Meteorological Society* **121**: 319–367.
Forsythe WC, Rykiel Jr EJ, Stahl RS, Wu H, Schoolfield RM. 1995. A model comparison for daylength as a function of latitude and day of year. *Ecological Modeling* **80**: 87–95.
Glahn HR, Chambers TL, Richardson WS, Perrotti HP. 1985. Objective map analysis for the local AFOS MOS program. NOAA Technical Memorandum NWS TDL 75. NOAA, US Department of Commerce.
Gleason E. 2002. Global daily climatology network, V1.0. National Climatic Data Center, 151 Patton Ave., Asheville, NC.
Higgins RW, Janowiak JE, Yao Y. 1996. A gridded hourly precipitation database for the United States (1963–1993). NCEP/CPC Atlas, No. 1. [Available for NCEP/CPC, W/NP52, Washington, DC 20 233].
Hu Q, Feng S. 2001. A southward migration of centennial-scale variations of drought/flood in eastern China and the western United States. *Journal of Climate* **14**: 1323–1328.
Hubbard KG. 2001. Multiple station quality control procedures. In *Proceedings of Automated Weather stations for Applications in Agriculture and Water Resources Management: Current Use and Future Perspectives*, Hubbard KG, Sivakumar MVK (eds), Lincoln, Nebraska, USA, 6–10 March 2000; 133–136.
Janowiak JE, Bell GD, Chelliah M. 1999. A gridded database of daily temperature maxima and minima for the conterminous United States: 1948–1993. NCEP/Climate Prediction Center Atlas 6. [Available from the Climate Prediction Center, World Weather Building, Rm. 605, 5200 Auth Rd, Camp Springs, MD 20 746].
Jones PD, Raper SCB, Bradley RS, Diaz HF, Kelly PM, Wigley TML. 1986. Northern Hemisphere surface air temperature variations: 1851–1984. *Journal of Climate and Applied Meteorology* **25**: 161–179.
Kaiser DP, Tao S, Fu C, Zeng Z, Zhang Q, Wang W, Karl T. 1993. Climate data bases of the People's Republic of China, 1841–1988. Department of Energy Technical Report, TR-055.
Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* **77**: 437–471.
Klein Tank AMG, Wijngaard JB, Können GP, Böhm R, Demarée G, Gocheva A, Mileta M, Pashiardis S, Hejkrlik L, Kern-Hansen C, Heino R, Bessemoulin P, Müller-Westermeier G, Tzanakou M, Szalai S, Pálsdóttir T, Fitzgerald D, Rubin S, Capaldo M, Maugeri M, Leitass A, Bukantis A, Aberfeld R, van Engelen AFV, Forland E, Mietus M, Coelho F, Mares C, Razuvaev V, Nieplova E, Cegnar T, Antonio López J, Dahlström B, Moberg A, Kirchhofer W, Ceylan A, Pachaliuk O, Alexander LV, Petrovic P. 2002. Daily dataset of 20th-century surface air temperature and precipitation series for European climate assessment. *International Journal of Climatology* **22**: 1441–1453.
Kubecka P. 2001. A possible world record maximum natural ground surface temperature. *Weather* **56**: 218–221.
Lanzante JR. 1996. Resistant, robust and nonparametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. *International Journal of Climatology* **16**: 1197–1226.
Lawrimore JH, Peterson TC. 2000. Pan evaporation trends in dry and humid regions of the United States. *Journal of Hydrometeorology* **1**: 543–546.
Manton MJ, Della-Marta PM, Haylock MR, Hennessy KJ, Nicholls N, Chambers LE, Collins DA, Daw G, Finet A, Gunawan D, Inape K, Isobe H, Kestin TS, Lefale P, Leyu CH, Lwin T, Maitrepierre L, Ouprasitwong N, Page CM, Pahalad J, Plummer N, Salinger MJ, Suppiah R, Tran VL, Trewin B, Tibig I, Yee D. 2001. Trends in extreme daily rainfall and temperature in Southeast Asia and the South Pacific: 1961–1998. *International Journal of Climatology* **21**: 269–284.
Meek DW, Hatfield JL. 1994. Data quality checking for single station meteorological databases. *Agricultural and Forest Meteorology* **69**: 85–109.
Peterson TC, Easterling DR. 1994. Creation of homogeneous composite climatological reference series. *International Journal of Climatology* **14**: 671–679.
Peterson TC, Dann H, Jones P. 1997. Initial selection of a GCOS surface network. *Bulletin of the American Meteorological Society* **78**: 2145–2152.
Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Torok S, Auer I, Boehm R, Gullett D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger J, Eirik J, Førland EJ, Hanssen-Bauer I, Alexandersson H, Jones P, Parker D. 1998. Homogeneity adjustments of *in situ* atmospheric climate data: a review. *International Journal of Climatology* **18**: 1493–1517.
Qian W, Zhu Y. 2001. Climate change in China from 1880 to 1998 and its impact on the environmental condition. *Climatic Change* **50**: 419–444.
Qian WH, Zhang HN, Zhu YF, Lee DK. 2001. Interannual and interdecadal variability of East Asian Acas and their impact on temperature of China in winter season for the last century. *Advances in Atmospheric Sciences* **18**: 511–523.
Reek T, Doty SR, Owen TW. 1992. A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network. *Bulletin of the American Meteorological Society* **73**: 753–762.
Shen D, Varis O. 2001. Climate change in China. *Ambio* **30**: 381–383.
Stooksbury DE, Idso CD, Hubbard KG. 1999. The effects of data gaps on the calculated monthly mean maximum and minimum temperatures in the continental United States: a spatial and temporal study. *Journal of Climate* **12**: 1524–1533.
Tao S, Fu C, Zeng Z, Zhang Q, Kaiser D. 1991. Two long-term instrumental climatic data bases of the People's Republic of China. Oak Ridge National Laboratory ORNL/CDIAC-47, Oak Ridge, TN.

Vincent LA, Zhang X, Bonsal BR, Hogg WD. 2002. Homogenization of daily temperatures over Canada. *Journal of Climate* **15**: 1322–1334.
Wang JXL, Gaffen DJ. 2001. Late-twentieth-century climatology and trends of surface humidity and temperature in China. *Journal of Climate* **14**: 2833–2845.
Wijngaard JB, Klein Tank AMG, Können GP. 2003. Homogeneity of 20th century European daily temperature and precipitation series. *International Journal of Climatology* **23**: 679–692.
Yan Z, Yang C, Jones PD. 2001. Influence of inhomogeneity on the estimation of mean and extreme temperature trends in Beijing and Shanghai. *Advances in Atmospheric Sciences* **18**: 309–322.

