Computers & Geosciences 54 (2013) 309–317

Contents lists available at SciVerse ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Computer assisted screening, correction, and analysis of historical weather measurements

Dorian J. Burnette [a,*], David W. Stahle [b]

[a] Department of Earth Sciences, 109 Johnson Hall, University of Memphis, Memphis, TN 38152, USA
[b] Department of Geosciences, 113 Ozark Hall, University of Arkansas, Fayetteville, AR 72701, USA

A R T I C L E I N F O

Article history:
Received 20 August 2012
Received in revised form
15 January 2013
Accepted 17 January 2013
Available online 25 January 2013

Keywords:
Historical climate data
Homogenization toolkit

A B S T R A C T

A computer program, Historical Observation Tools (HOB Tools), has been developed to facilitate many of the calculations used by historical climatologists to develop instrumental and documentary temperature and precipitation datasets and makes them readily accessible to other researchers. The primitive methodology used by the early weather observers makes the application of standard techniques difficult. HOB Tools provides a step-by-step framework to visually and statistically assess, adjust, and reconstruct historical temperature and precipitation datasets. These routines include the ability to check for undocumented discontinuities, adjust temperature data for poor thermometer exposures and diurnal averaging, and assess and adjust daily precipitation data for undercount. This paper provides an overview of the Visual Basic.NET program and a demonstration of how it can assist in the development of extended temperature and precipitation datasets using modern and early instrumental measurements from the United States.

1. Introduction

Historical weather observations were recorded using nonstandard methodologies, and therefore require careful screening and correction routines to extract the highest quality weather data that are also compatible with modern observations (Chenoweth, 1992, 1993; Daly et al., 2007). Unfortunately, changes in the observing methods can occur at irregular intervals and increase the difficulty in developing and utilizing fully automated numerical techniques. Nevertheless, automated screening and correction can proceed in a segmented step-wise fashion under direct control by the analyst. The large quantity of data available for the analysis and recovery of historical weather measurements makes the use of spreadsheets inefficient and risks the introduction of additional error. Therefore, there is a need to automate these screening and correction procedures with computerized programs. The dendrochronology (tree-ring dating) community has used computerized correlation analyses for quality control in tree-ring dating and to identify problematic series and segments (COFECHA; Holmes, 1983; Grissino-Mayer, 2001). The computer program ARSTAN is also widely used to standardize tree-ring data and to develop the mean chronologies (Cook, 1985). The purpose of this research has been to develop a computer-assisted analytical framework for the assessment and recovery of instrumental daily temperature and precipitation measurements from the early decades of meteorological observation.

Much of the recovery and development of long, daily instrumental temperature and precipitation records has occurred in Western Europe (e.g., Camuffo and Jones, 2002; Slonosky, 2003). Venema et al. (2012) assessed key algorithms used in the development of homogeneous weather records, but these algorithms only work with monthly data. Almost no algorithms exist to aid in the homogenization of daily historical weather records, and those that do require a homogenous reference series for comparison, where a candidate station is examined relative to a neighboring station determined to be free of discontinuities (e.g., Alexandersson, 1986; Mestre et al., 2011). Historical weather records across the United States, such as those available from the U.S. Army Surgeon General, the Smithsonian, and the U.S. Signal Service, have not been widely developed and are likely to have issues with undocumented discontinuities and exposure (Chenoweth, 1992, 1993). Therefore, homogeneous reference series have to be developed from scratch. Here we develop a set of computer programs, Historical Observation Tools (HOB Tools), in Microsoft Visual Basic.NET to assist in the recovery and development of homogeneous daily, monthly, seasonal, and annual temperature and precipitation data across the United States. We demonstrate HOB Tools with daily instrumental weather measurements taken at stations in Kansas during the 19th century (Burnette et al., 2010; Burnette and Stahle, 2013), with the term "tools" referring to the various computation routines included in the system. The tools can be accessed via the main menu system located at the top of the program (Fig. 1).

* Corresponding author. Tel.: +1 901 678 4452; fax: +1 901 678 2178.
E-mail address: djbrntte@memphis.edu (D.J. Burnette).

0098-3004/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cageo.2013.01.003

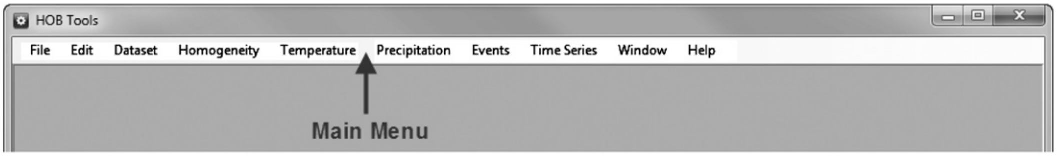

Fig. 1. The HOB Tools graphical user interface is run from the main menu system located at the top of the program. Each sub-menu contains a group of tools. There are seven tools available to format and quality control the temperature and precipitation data, so they can be used in the 20 other functional analyses, which screen, adjust, reconstruct, and plot historical temperature and precipitation data.

HOB Tools requires Windows 2000 or later and Microsoft.NET Framework 2.0 or higher. The program is available as a compiled executable file with its own setup program. Thus, users do not need to manipulate the source code in order to run datasets through the program. HOB Tools comes with a user's guide that (1) describes how to run the program, (2) provides a general description of each computation tool, and (3) provides the input information required to operate each tool. The program reads and produces data in tab-delimited ASCII format, except for the data downloaded from the U.S. Historical Climatology Network (USHCN), Global Historical Climatology Network (GHCN), or Integrated Global Radiosonde Archive (IGRA) websites which are read in their native formats. The tab-delimited format is used because it is the most convenient when digitizing written historical measurements. Most HOB Tools operations output data to tab-delimited ASCII files, and eight tools create graphical data for visual inspection. All numerical and statistical calculations performed by HOB Tools were tested and verified in the statistical package R (R Core Development Team, 2012).

The current version of HOB Tools represents a first step toward synthesizing the helpful routines used in the recovery and reanalysis of historical daily temperature and precipitation measurements. HOB Tools is an open source project, licensed under the GNU General Public License, and available online at http://www.djburnette.com/programs/hobtools/. Users who would like to contribute to the project are welcome to download the source code and edit existing or add new tools to the program (see "readme.txt" included in the source code for further instructions). The goals of this open source project are (1) to allow for the free scientific exchange of ideas that could improve the methods of historical weather and climate analysis and (2) to make these methods readily accessible to the research community.

2. Quality control of newly digitized datasets

Early instrumental data must often be digitized from faded documents or microfilms for historical climate research. Extreme care must be exercised when these data are digitized, so no additional error is introduced (e.g., a day's worth of data could be missed by accident). Newly digitized station data are first submitted to the "Data Entry and Gap Quality Control" tool, which assumes that the time series sequence for input data is continuous and any missing time periods or back-to-back days when the same data were digitized are listed in the output. Any time periods that were not associated with known gaps should be checked to verify that the data were digitized correctly.

3. Temperature reconstruction and analysis

A series of tools are available in the HOB Tools "Temperature" sub-menu to assist in the recovery of homogeneous temperature data from historical sources (Fig. 1). All of these tools were used to develop the daily mean temperature reconstruction for Manhattan, Kansas (Burnette et al., 2010). Computational tools in the "Dataset," "Homogeneity," "Events," and "Time Series" sub-menus were also required during the development and analysis of temperature data for eastern Kansas and may be helpful.

3.1. Exposure changes and other undocumented discontinuities

Historical temperature observations often contain problems with the exposure of the thermometer due to the non-standard practices of early observers. In fact, there can be undocumented changes in the thermometer exposure or station location at any time in the history of a station. The Standard Normal Homogeneity Test (SNHT) developed by Alexandersson (1986) has widely been used to detect discontinuities. However, a homogeneous reference series is required for this test. Unfortunately, neighboring 19th century stations in the U.S. have not been assessed for discontinuities, and greatly limit the confidence in the results obtained from the SNHT. Therefore, another method is needed to adequately detect discontinuities in the 19th century data from the U.S.

The "Screen Interhourly Temperature Ranges" tool was used to detect changing exposure or potential changes in observation methodology (i.e., "discontinuities" in the station record). The tool performs a visual and statistical analysis of the diurnal temperature range between morning and afternoon observations in the historical period relative to those in the modern temperature record. All modern observations can be extrapolated to the top of the hour, solar time, and sunrise/sunset as required so they are compatible as possible with historical measurements. Time series plots of the historical diurnal temperature range are then displayed (e.g., Fig. 2). July is displayed by default because abrupt excursions in the diurnal temperature range time series can be detected more easily in the summer due to the smaller temperature variance. Any abrupt shift in the time series is likely to be a discontinuity (e.g., 1883/84 in Fig. 2). In fact, metadata indicate that the temperature observations at Manhattan were recorded by a different observer at a different location within the city beginning in 1884. This change was well-documented on the U.S. Signal Service meteorological registers, and demonstrates how the diurnal temperature range can be helpful with the detection of undocumented discontinuities.

Visual inspection of the diurnal temperature range can be inconclusive if the historical station record is short, and the precise time when a discontinuity is introduced can be masked by the large variability in temperature during the winter season. Thus, this tool also flags daily temperature ranges in the historical period that are below the 5th and above the 95th modern percentiles. An abrupt change in the occurrence of these flags could suggest a potential discontinuity in the historical measurements.

3.2. Homogenization of segmented station data

Burnette et al. (2010) separated all daily instrumental measurement time series for every station at each documented or detected discontinuity. Each station segment was then corrected for poor thermometer signals using the "Adjust Station Segments"

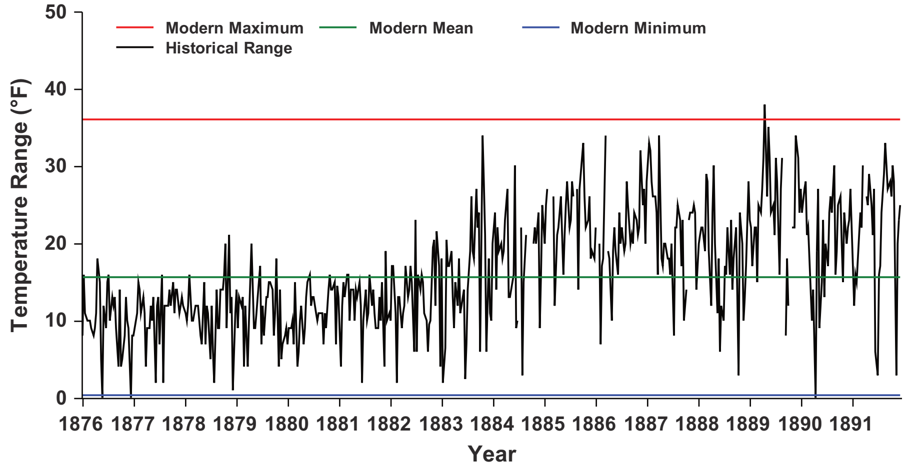

Fig. 2. The daily range in temperature from 6 a.m. to 1 p.m. at Manhattan, Kansas, during July from 1876 to 1892 is illustrated. Maximum, minimum, and mean ranges between the same 2 h in the modern record are also plotted, so the historical data can be visually assessed relative to the modern. A discontinuity can be seen by the change in mean and variance of the time series at 1883/84.

tool. This tool adjusts all fixed hourly or minimum and maximum temperatures relative to the modern diurnal temperature cycle (e.g., Chenoweth, 1998; Bergström and Moberg, 2002). Historical and modern temperature data are transformed into anomalies relative to their individual time periods. Differences between the fixed hourly or minimum and maximum observations and the same observations in the modern period are then calculated on a monthly basis. These differences yield monthly correction values that are then applied to the input historical data.

The output from the "Adjust Station Segments" analysis consists of two files. The first contains the adjusted fixed hourly or minimum and maximum temperature data. The second file quantifies the actual adjustments that were performed on the input data. The results from this second output file were analyzed by Burnette et al. (2010) for potential clues about the exposure of the thermometer. HOB Tools also performs a quality control check on the temperature data when this tool is run, and a warning message is generated if any temperature value is $\geq 60\,^{\circ}$C or $\leq -40\,^{\circ}$C. This allows for detection of temperature data that were incorrectly recorded or digitized (e.g., a value of 533 was entered instead of 53).

3.3. Derivation of daily mean temperature

Daily mean temperatures can be computed for each homogenized station segment using the "Compute Daily Mean Temperatures" tool. If daily means are computed from minimum and maximum temperatures, then Eq. (1) is used:

$$T_{\text{mean}} = (T_{\text{min}} + T_{\text{max}})(2)^{-1} \tag{1}$$

where T_{mean} is the daily mean temperature, T_{min} is the minimum temperature, and T_{max} is the maximum temperature.

If daily means are computed from fixed hourly temperature readings (e.g., 7 a.m., 2 p.m., and 9 p.m.), then modern hourly temperature data are required, which allows HOB Tools to perform corrections for diurnal averaging as discussed in Burnette et al. (2010). The requirement of modern hourly data could limit the usefulness of this method in some areas, but such data have been used to correct daily temperature data outside of the U.S. (e.g., Konnen et al., 2003).

Daily mean temperatures are computed from fixed hourly temperatures by calculating an average difference between a daily mean derived using Eq. (1) and a daily mean derived by averaging the fixed hourly temperatures in the relevant calendar month. The difference is then scaled according to the ratio of the range between morning and afternoon temperatures on the target historical day, R_{hf}, and the modern average range from the equivalent fixed hourly observations in the same calendar month (R_{mf}):

$$C = (M_{\text{mm}} - M_{\text{mf}})(R_{\text{hf}})(R_{\text{mf}})^{-1} \tag{2}$$

where C is the correction for diurnal averaging, M_{mm} is the average modern daily mean temperature from maximum and minimum values and M_{mf} is the average modern daily mean temperature from fixed hourly observations.

For example, if the average of maximum and minimum temperature in April was 13 °C and the average of 7 a.m., 2 p.m., and 9 p.m. temperatures in the same month was 14 °C, then the difference between M_{mm} and M_{mf} would be -1 °C. This is the value that would then be used to transform a daily mean computed from the fixed hourly readings common in the historical period into a daily mean computed from maximum and minimum temperatures common today. It accounts for the normal diurnal cycle in the given calendar month. The value, -1 °C, would then be scaled for day-to-day variations in the diurnal range arising from weather variations in the historical record (e.g., cloudiness). Therefore, if the average difference between the 7 a.m. and 2 p.m. temperatures in the modern record was 13 °C and the difference between the same 2 h on a given day in the historical period was only 1 °C due to cloudy conditions, then R_{hf} would be 1 °C and R_{mf} would be 13 °C. The final correction, C, would then be $(-1)(1)(13)^{-1}$ or approximately -0.08 °C. In this case, the cloudy conditions result in little correction. This accounting for day-to-day weather variations does not impact the overall reconstruction much, but it can yield important differences on the order of 1 °C or more for individual days (e.g., Bergström and Moberg, 2002).

The correction, C, is then applied to the daily mean temperature, which is calculated as:

$$T_{\text{mean}} = (T_{\text{am}} + T_{\text{pm}} + T_{\text{eve}})(3)^{-1} + C \tag{3}$$

where T_{am} is the morning, T_{pm} is the afternoon, T_{eve} is the evening temperature, and C is the correction value that adjusts the historical daily means for compatibility with modern daily means. Again, the program will validate that each fixed hourly or minimum and maximum temperature value falls between -40 °C and 60 °C.

3.4. Evaluation of the station segments

Ideally, all station data should be separated at each documented or detected discontinuity and treated as separate records. However, it is possible that in light of the adjustments applied in Sections 3.2 and 3.3, some discontinuities may no longer be detectable beyond the background of climate variability. There are two tools under the "Homogeneity" sub-menu that can be useful for such assessments.

One way to assess potential discontinuities within one station record is to compare the data with another station that has no known discontinuities during the time period in question. The "Difference Series Plot" makes this comparison over two different segments of time that are entered by the user and assumes that a discontinuity occurred between Segments 1 and 2 (Fig. 3). Any statistically significant difference or other spurious behavior in the difference series would suggest a discontinuity in the suspect station. However, short-term noise introduced naturally by the movement of frontal systems can impact this analysis. Thus, these difference series analyses can also be performed at monthly, seasonal, and annual resolution to help isolate true problem segments.

A simple graphical analysis of stations relative to one another can also assist in the identification of problematic segments. The "Spaghetti Plot" tool illustrates station data between user-specified start and end dates (inclusive), and can highlight a single station of interest (e.g., Fig. 4). This highlighted station can then be visually assessed relative to the rest of the network. The plotted time series can be inspected closely using the zooming and panning capabilities of the plots in HOB Tools. While this tool may be subjective, a station with problematic temperature data will diverge from the rest of the network (e.g., Fig. 4), and such data must be removed before the reconstruction proceeds.

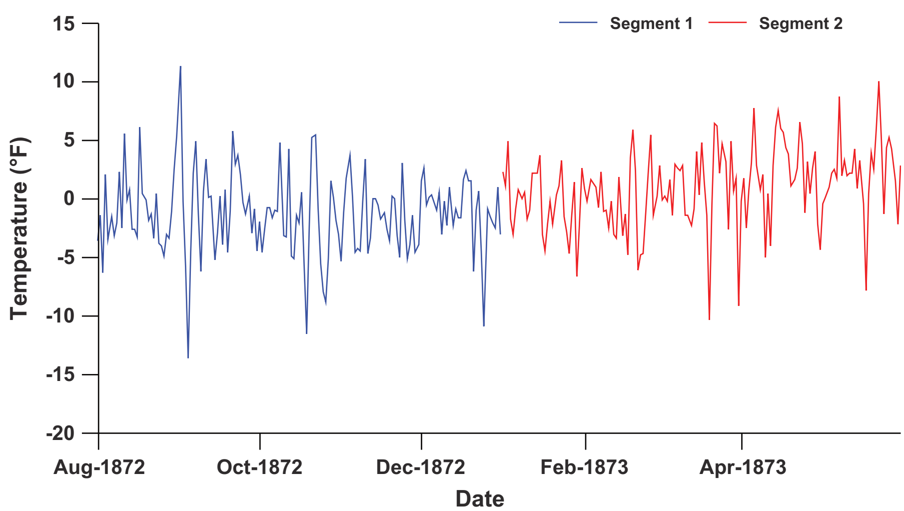

Fig. 3. A potential discontinuity occurred in the daily mean temperature record at Leavenworth, Kansas, on 1 January 1873 (Burnette et al., 2010). Visual inspection of difference series between daily mean temperatures at Manhattan and Leavenworth before and after the discontinuity suggests a large change in the mean (blue series vs. red series). Note the trend in the time series during the spring of 1873 and the lack of trend from August through December 1872. HOB Tools performed a two-tailed *t*-test that suggested the discontinuity was statistically significant ($p < 0.0001$, not shown).

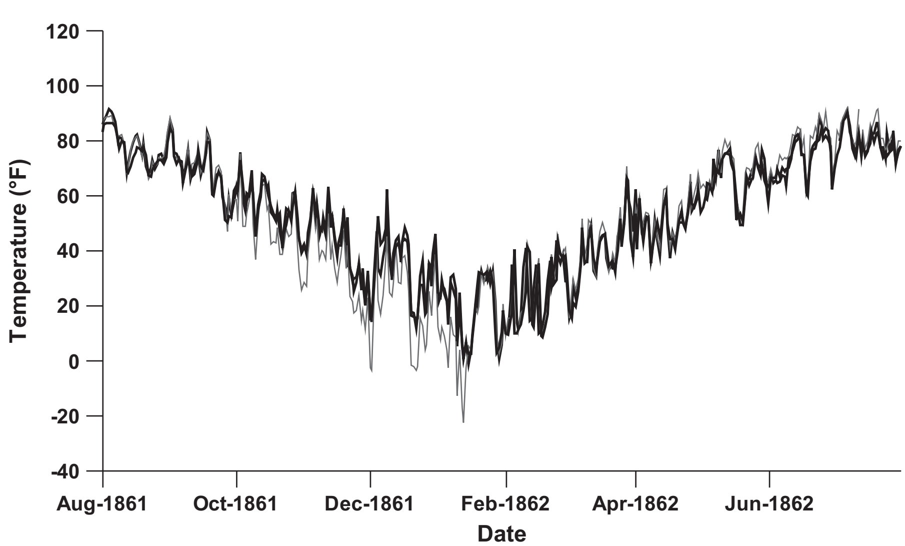

Fig. 4. Daily mean temperature from a selected number of stations in eastern Kansas were generally in agreement with one another from August 1861 through July 1862 (black time series). The lone exception was Fort Riley, Kansas (gray time series). Note the deviation of the Fort Riley record in late 1861, and its greater consistency with the other regional stations in 1862.

3.5. Reconstructing daily temperature from discontinuous station segments

The ''Reconstruct Temperature'' tool is used for the reconstruction of daily temperature data from several homogenous stations or station segments. The reconstruction process is not yet fully computerized in HOB Tools because operator intervention is often required to choose the optimal data and calibration models for the longest possible estimates of daily temperature. It should be possible to automate many of these decisions using restrictive thresholds for model statistics, but this will inevitably lead to shorter and more discontinuous historical reconstructions. Future versions of HOB Tools are planned that will include both manual step-by-step and automated temperature reconstruction options.

The reconstruction of daily mean temperature is described in Burnette et al. (2010), and the procedure used in HOB Tools is detailed in the user's guide and summarized below. First, the user identifies the dependent station that will be reconstructed and the independent surrounding stations that will be used to fill gaps in the dependent station and extend it backward in time. The independent station data are then transferred to the dependent station via simple linear regression separately, and the tool pauses to allow the user to assess the quality of the regression models and analyze plots of the residuals. Large outlying residuals, low R^2, high standard error, and a low coefficient of efficiency (CE) may together suggest that problematic daily data are degrading the relationship between the two stations (Draper and Smith, 1981; Cook et al.,1994; Kutner et al., 2005; Wilks, 2006). The user can select the outlying residuals that require further inspection, and those station data will then be assessed relative to the rest of the network with the most outlying station data removed from the regression modeling routine [see Burnette et al. (2010) or the HOB Tools user's guide for further details].

Final output from the ''Reconstruct Temperature'' tool consists of three files. The main output file contains the daily temperature reconstruction along with the standard error of each estimate and the name of the station used. Another file, with the words ''Test Models'' appended to the end of the filename, contains the forward and backward calibration and verification results for each station separated by a dashed line. The third file, with the words ''Final Models'' appended to the end of the name, contains the models that were used to transfer daily temperature data from each independent station to the dependent station. Each station is separated by a dashed line.

3.6. Modern data

Modern station data can be downloaded from the GHCN and USHCN websites. Direct links to these data repositories and download instructions are available within the HOB Tools user's guide. The temperature and precipitation data available are in a different ASCII file format though, and require conversion after they are downloaded. There are three tools within the ''Dataset'' sub-menu, which transform these data into the tab-delimited ASCII format used by HOB Tools.

3.7. Reconstruction and homogenization of the modern data

Homogenization of modern temperature records in the United States has been done only at monthly resolution based on known changes in the time temperatures were observed and change-point detection procedures that identify other changes in observation practice (e.g., USHCN; Williams et al., 2006; Menne et al., 2009). Thus, modern daily data can still contain various non-climatic biases that should be minimized. Previous studies have adjusted daily temperature data relative to a homogeneous reference series (cf., Jones and Lister, 2002; Moberg et al., 2002; Vincent et al., 2002). The ''Adjust Daily Temperature Data'' tool follows this methodology, where monthly averages of the daily temperature data are calculated and compared with the monthly data from the homogenous reference series. This comparison yields a correction value that can then either be applied directly to the daily data or the adjustments can be performed by interpolating between the monthly correction values. The latter is recommended given the possibility that additional discontinuities could be introduced using the former method. All monthly correction values and the adjusted daily data are output. The availability of a ''homogeneous reference series'' is of course crucial to this analysis. In the modern records after 1894, homogenized monthly temperature data are available from the USHCN and the GHCN. Prior to 1895, ''homogenous reference series'' are much more difficult to acquire, and HOB Tools is designed to screen, reconstruct, and assess the highest quality historical daily records.

3.8. Temperature analysis

Ideally, all screening and correction of the historical data should be performed within its original units and then later converted to units suitable for publication. Unfortunately, conversion between different units of temperature may be required due to the nature of the input data. Thus, some programs within HOB Tools have a unit conversion capability built in (e.g., ''Compute Daily Mean Temperatures''). For a quick conversion of an entire dataset, the ''Convert Temperature Units'' tool can be used to transform input temperature data from one set of temperature units into another.

The daily mean temperature data reconstructed for Manhattan, Kansas (Burnette et al., 2010), were run through various analyses using the ''Temperature Analysis'' tool, which creates time series of mean, minimum, and maximum temperatures and heating/cooling degree days. Heating (*HDD*) and cooling (*CDD*) degree days are simply a summation of daily mean temperature (T_{mean}) from a base temperature (T_{base}) input by the user. The default value for T_{base} is 18 (i.e., the heating and cooling degree day base temperature in °C):

$$HDD = \sum_{\text{Annual}} (T_{base} - T_{mean})$$

$$CDD = \sum_{\text{Annual}} (T_{mean} - T_{base}) \quad (4)$$

The temperature analyses can be run during specific months or during a specific range of months (e.g., seasonal, annual). Any missing data encountered by HOB Tools results in a flag (including values needed to compute error bars). The number of flags is reported in the output file, and any result other than zero should be checked to determine if the missing data impact the results.

Other important analysis tools are found under the ''Events'' and ''Time Series'' sub-menus. One tool, ''Cold Wave Analysis,'' calculates the frequency of cold waves from daily temperature data based on a daily temperature depression input by the user. The analysis can be performed per month or across a range months (e.g., during the winter season). A ''Return Time Analysis'' tool is also available, which calculates the probability that a specific event or critical value will occur again over a specified number of years. The frequency a user-specified threshold value is met can be determined using the ''Threshold Analysis'' tool. The threshold value can be a specific value or a percentile. Finally, there is a basic ''Time Series Plot'' tool, which produces simple plots of time series in monthly, seasonal, or annual format and generates summary statistics.

4. Precipitation reconstruction and analysis

The ‘‘Precipitation’’ sub-menu system in HOB Tools follows the general procedure used to develop the spring, summer, and growing season precipitation reconstructions for northeastern Kansas and northwestern Missouri (Burnette and Stahle, 2013). The same tools also include the ability to analyze and reconstruct cold season precipitation, which can contain serious undercount bias (Larson and Peck, 1974). Any modern precipitation downloaded from the National Climatic Data Center or elsewhere will typically need to be reformatted using the same tools within the ‘‘Dataset’’ sub-menu. Other tools listed under the ‘‘Homogeneity,’’ ‘‘Events,’’ and ‘‘Time Series’’ will also be required during the development and analysis of historical precipitation data.

4.1. The frequency distribution of daily precipitation amounts

Precipitation data can be more difficult to correct than the temperature data, so it is helpful to focus on the stations with the highest quality observations. The ‘‘Screen Precipitation Data’’ tool computes the frequency distribution of daily precipitation amounts. Precipitation frequency distributions are created for every hundredth of an inch from 0.01 to 2.00 in.. The results for every day and for each month are output separately to the output file and displayed in histograms (e.g., Fig. 5). These histograms of daily precipitation frequency can also be plotted for specific months or seasons. A daily precipitation frequency histogram arranged in order from frequent low amounts on the left side to infrequent high amounts on the right side should exhibit a smooth negative exponential curve with no significant truncation of low amounts or stair-step appearance if the daily precipitation data are of high-quality (e.g., Fig. 5b; Daly et al., 2007).

4.2. Compute monthly totals

The best stations with daily precipitation data were then run through the ‘‘Compute Precipitation Totals’’ tool, which sums daily or monthly precipitation amounts and precipitation days for each station. These analyses can be run during specific months or for a specific range of months (e.g., seasonal, annual). Given that undercount is a common bias in historical precipitation data (Larson and Peck, 1974; Mock, 1991, 2000), and analysis of thresholds can assist in quantifying such bias (Burnette and Stahle, 2013), HOB Tools can sum daily precipitation amounts and days above a certain threshold value. This tool also allows

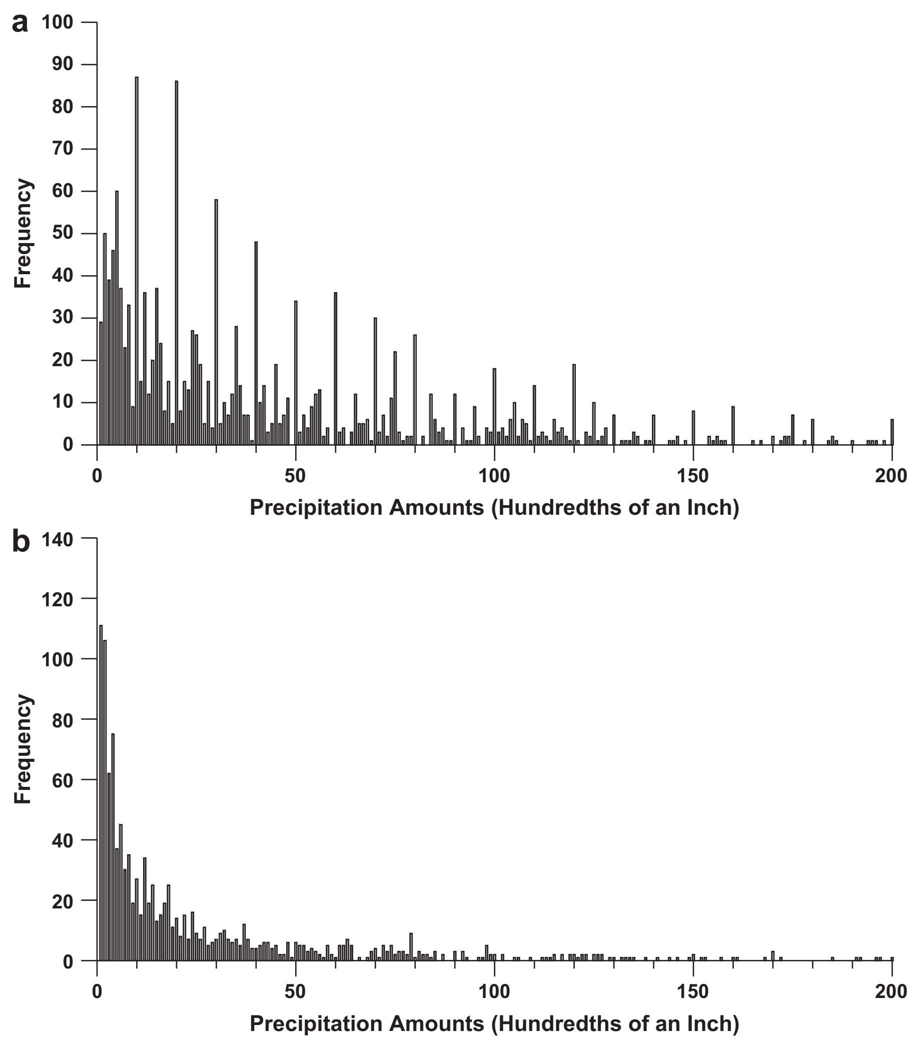

Fig. 5. (a) The frequency distribution of daily precipitation totals compute for each 0.01 in. at Fort Leavenworth, Kansas, for the 28-year period from 1855 to 1883 suggests that the observers may have been biased toward recording precipitation amounts to the nearest tenth of an inch and undercounting low daily totals, especially totals below 0.10 in.. (b) The frequency distribution of daily precipitation totals at Dodge City, Kansas, indicates that these data are more homogenous than the Fort Leavenworth data given the dominance of low precipitation totals, the smooth histogram, and the absence of obvious stair-step appearance (i.e., no indication observers were biased to recording precipitation data to the nearest tenth of an inch).

a special value, "-1", to denote a day when more than a trace of precipitation occurred but no amount was reported. This method accounts for observers who did not record a precipitation amount until the end of an event (e.g., the observer notes the rain started on the 13th, ended on the 14th, and recorded a 2-day total on the 14th). It also accounts for observers who only reported whether or not precipitation occurred. Both situations can be used for rain day summaries, but not for daily precipitation totals. Any missing data encountered by HOB Tools results in a flag (including data missing from the calculations of error bars). The number of flags is reported in the output file, and flags should be checked to determine if the missing data impact the results.

4.3. Estimate of unrecorded light precipitation amounts

To minimize bias arising from the failure to record light precipitation amounts in certain early instrumental records, a threshold procedure may be used to sum precipitation amounts (or days) above a specified threshold (e.g., ≥ 0.10 in.). This threshold series can be computed using the "Compute Precipitation Totals" tool (Section 4.2), and then input with the raw precipitation totals into the "Estimate Precipitation" tool. This tool compares a reconstruction of precipitation based on the total amount with the amount that exceeds a specified threshold during a user-defined period when the data are presumed to be high quality. An estimate of the average precipitation amount typically observed below the threshold during the high quality data period is then added to the totals above the threshold input by the user during the time period suspected to suffer from undercount. This method to estimate the true total amount of precipitation can be performed at monthly, seasonal, or annual resolution, but it assumes that the observer(s) always failed to record precipitation amounts below the threshold.

When the "Estimate Precipitation" tool is run, the estimation method is tested by splitting the period of high quality data in half. An estimate of the trace precipitation amounts below the threshold is then computed over the last half of the high quality period and the results are output under "Initial Derived Differences". The estimates are then added to the threshold series during the first half of the high quality period. The new estimated precipitation amounts are then compared with the actual precipitation amounts and summary statistics are output under "Verification of Initial Derived Differences". An estimate of the trace precipitation amounts below the threshold is then computed over the entire high quality period and the results are output under "Final Derived Differences". Low average errors under "Verification of Initial Derived Differences" and similar values between the "Initial Derived Differences" and "Final Derived Differences" are ideal. All of these results are reported in a separate output file with the word "Statistics" appended to the end of the filename. The tool then adds the "Final Derived Differences" to the threshold data time series to create an adjusted precipitation time series during the user-defined period of adjustment. If monthly data are input, then each month is processed and estimated separately. Once the end of the adjustment period is reached, HOB Tools appends the remaining raw precipitation time series to the end of the newly adjusted dataset.

4.4. Reconstruct precipitation

Given the spatial variability of precipitation, especially during the warm season, it may be helpful to combine the highest quality precipitation data from multiple stations into a regional average of monthly or seasonal total precipitation (e.g., Burnette and Stahle, 2013). The "Reconstruct Precipitation" tool develops a regional-average of precipitation amounts or days on a monthly basis by averaging a series of stations input by the user. The interquartile range of the precipitation observations is used to approximate the amount of uncertainty associated with the reconstructed value. A change in the number of stations used to develop the precipitation reconstruction could induce spurious variability in the reconstructed time series. Thus, this tool also includes the capability to always use the same number of stations throughout the reconstruction. If that option is selected, then HOB Tools will calculate the precipitation reconstruction by only using the user-specified number of stations closest to the network average. An option that uses a weighting scheme to determine a regional average would be a useful addition to the program in a future release to minimize issues with spatial clustering.

4.5. Precipitation homogeneity

The paucity of historical weather records within close proximity to one another and the localized nature of warm season convective precipitation make the assessment of homogeneity in the reconstructed precipitation estimates difficult. The "Spaghetti Plot" tool can be used to visually examine all the precipitation data relative to one another, which may identify outliers. Another tool, "Double Mass Analysis," can be used to assess relative homogeneity of two precipitation series (Kohler, 1949). This tool, inspired by the Interactive Double Mass Analysis program used by the Office of Hydrological Development at the National Oceanographic and Atmospheric Administration (NOAA, 1999), sums data from two datasets, generates a double mass plot, and then uses simple linear regression and user-specified input information to detect breakpoints or analyze a suspected breakpoint in the cumulative data. All breakpoints are specified in the text output file and are indicated on the double mass plot.

4.6. Precipitation analysis

Precipitation totals in (Burnette and Stahle, 2013) were assessed and developed in units of inches. These data were then converted into millimeters using the "Convert Precipitation Units" tool. Like temperature data, final precipitation reconstructions can be further analyzed using tools under the "Events" and "Time Series" submenus. The probability that a specific seasonal precipitation amount or day count would occur again over a specified number of years can be analyzed with the "Return Time Analysis" tool. The frequency with which a user-specified threshold value is met can be determined using the "Threshold Analysis" tool. Finally, the "Time Series Plot" tool can be used to produce summary statistics and plots of precipitation totals or day counts.

5. Conclusions

Rigorous screening and correction routines are imperative to maximize the value of historical weather observations. A toolkit of menu-driven computer programs, Historical Observation Tools (HOB Tools), has been developed that compiles many of the routines used for the screening, recovery, and analysis of instrumental daily temperature and precipitation observations. The various changes in exposure, instrumentation, and station location that typically complicate historical temperature measurements often require that numerical processing be performed in a series of steps that can each be scrutinized before continuing to the next procedure. As currently configured, HOB Tools requires the user to run through a set of procedures (i.e., "tools") to analyze the data quality visually and statistically in order to obtain a final temperature (or precipitation) reconstruction. The reconstruction of daily mean temperature from 1 July 1828 to 28

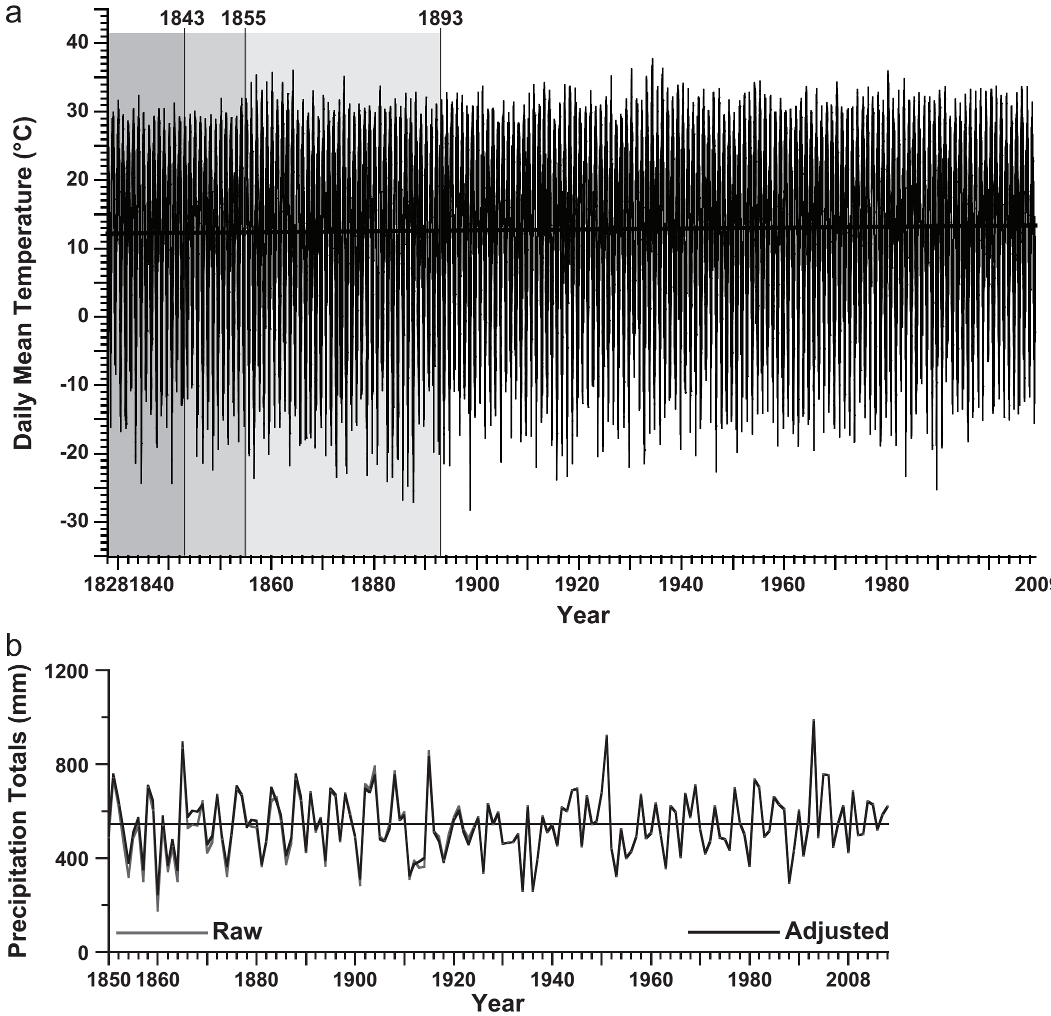

Fig. 6. (a) The temperature reconstruction for eastern Kansas includes 65,987 daily mean estimates and dates from 1 July 1828 to 28 February 2009 (Burnette et al., 2010). This is the longest unbroken daily mean temperature record yet published in the Americas, but the uncertainty of the estimates increases prior to 1893, 1855, and 1843. (b) Growing season (April through August) reconstruction of precipitation totals for northeastern Kansas and northwestern Missouri from 1850–2008 (Burnette and Stahle, 2013). Note the raw precipitation totals prior to 1925 and the adjusted precipitation estimates that were computed using the "Estimate Precipitation" tool.

February 2009 for eastern Kansas (Burnette et al., 2010) and growing season precipitation totals for northwestern Missouri and northeastern Kansas (Burnette and Stahle, 2013) are illustrated in Fig. 6 and were both computed with the assistance of HOB Tools.

Development of HOB Tools has primarily centered on daily temperature and precipitation observations recorded in the United States by the Army Surgeon General, the Smithsonian Institution, the Signal Service, the National Weather Service, newspapers, and diarists with an interest in the weather. Additional observations from historical sources such as cloud classification have been valuable in the development of daily temperature datasets (e.g., Bergström and Moberg, 2002). Barometric pressure data also exist, which would be helpful for the reconstruction of synoptic-scale weather events (e.g., winter storms). However, further routines to correct the pressure data to sea-level and to 0 °C remain to be developed and added to HOB Tools. Thus, the current state of HOB Tools represents only the first step toward synthesizing the helpful routines for the recovery and reanalysis of historical climate records.

Acknowledgments

HOB Tools is an open source project, licensed under the GNU General Public License. The compiled executable file and all source code are freely available at http://www.djburnette.com/programs/hobtools/. We thank Rebecca Rowland, Matt Taylor, and Brad Johnson for providing feedback and bug reports that have improved this program. We also thank an anonymous reviewer for comments that have improved this manuscript. This research was sponsored by the National Science Foundation (DDRI Grant BCS-0622894).

Appendix A. Supplementary materials

Supplementary materials associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.cageo.2013.01.003.

References

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

Bergström, H., Moberg, A., 2002. Daily air temperature and pressure series for Uppsala (1722–1998). Climatic Change 53, 213–252.

Burnette, D.J., Stahle, D.W., 2013. Historical perspective on the Dust Bowl drought in the central United States. Climatic Change, doi:10.1007/s10584-012-0525-2.

Burnette, D.J., Stahle, D.W., Mock, C.J., 2010. Daily mean temperature reconstructed for Kansas from early instrumental and modern observations. Journal of Climate 23, 1308–1333.

Camuffo, D., Jones, P., 2002. Improved understanding of past climatic variability from early daily European instrumental sources. Climatic Change 53, 1–4.

Chenoweth, M., 1992. A possible discontinuity in the U.S. historical climate record. Journal of Climate 5, 1172–1179.

Chenoweth, M., 1993. Nonstandard thermometer exposures at U.S. cooperative weather stations during the late nineteenth century. Journal of Climate 6, 1787–1797.

Chenoweth, M., 1998. The early 19th century climate of the Bahamas and a comparison with 20th century averages. Climatic Change 40, 577–603.

Cook, E.R., 1985. A time series approach to tree ring standardization. Ph.D. dissertation, University of Arizona, Tucson, Arizona, USA, 171 pp. [Program ARSTAN available online at 〈http://www.ldeo.columbia.edu/res/fac/trl/public/publicSoftware.html〉.].

Cook, E.R., Briffa, K.R., Jones, P.D., 1994. Spatial regression methods in dendroclimatology: a review and comparison of two techniques. International Journal of Climatology 14, 379–402.

Daly, C., Gibson, W.P., Taylor, G.H., Doggett, M.K., Smith, J.I., 2007. Observer bias in daily precipitation measurements at United States Cooperative Network Stations. Bulletin of the American Meteorological Society 88, 899–912.

Draper, N.R., Smith, H., 1981. Applied Regression Analysis, 2nd ed. John Wiley 709 pp..

Grissino-Mayer, H.D., 2001. Evaluating crossdating accuracy: a manual and tutorial for the computer program COFECHA. Tree-Ring Research 57, 205–221.

Holmes, R.L., 1983. Computer assisted quality control in tree-ring dating and measurement. Tree-Ring Bulletin 43, 69–78.

Jones, P.D., Lister, D.H., 2002. The daily temperature record for St. Petersburg. Climatic Change 53, 253–267.

Kohler, M., 1949. On the use of double-mass analysis for testing the consistency of meteorological records and for making required adjustments. Bulletin of the American Meteorological Society 30, 188–189.

Konnen, G.P., Zaiki, M., Baede, A.P.M., Mikami, T., Jones, P.D., Tsukahara, T., 2003. Pre-1972 extension of the Japanese instrumental meteorological observation series back to 1819. Journal of Climate 16, 118–131.

Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W., 2005. Applied Linear Statistical Models, 5th ed. McGraw-Hill, New York, New York, USA 1396 pp..

Larson, L.W., Peck, E.L., 1974. Accuracy of precipitation measurements for hydrologic modeling. Water Resources Research 10, 857–863.

Menne, M.J., Williams Jr., C.N., Vose, R.S., 2009. The United States historical climatology network monthly temperature data—version 2. Bulletin of the American Meteorological Society 90, 993–1007.

Mestre, O., Gruber, C., Prieur, C., Caussinus, H., Jourdain, S., 2011. SPLIDHOM: a method for homogenization of daily temperature observations. Journal of Applied Meteorology and Climatology 50, 2343–2358.

Moberg, A., Bergström, H., Ruiz Krigsman, J., Svanered, O., 2002. Daily air temperature and pressure series for Stockholm (1756–1998). Climatic Change 53, 171–212.

Mock, C.J., 1991. Drought and precipitation fluctuations in the Great Plains during the late nineteenth century. Great Plains Research 1, 26–57.

Mock, C.J., 2000. Rainfall in the garden of the United States Great Plains, 1870–1889. Climatic Change 44, 173–195.

National Oceanic and Atmospheric Administration, 1999. Interactive Double Mass Analysis User's Guide, Version 1.2, 〈http://www.nws.noaa.gov/ohd/hrl/idma/users_guide/html/dma_overview.php〉. [accessed on 17.05.12].

R Core Development Team, 2012. The R Project for Statistical Computing, 〈http://www.r-project.org〉. [accessed on 17.05.12].

Slonosky, V.C., 2003. The meteorological observations of Jean-Francois Gaultier, Quebec, Canata 1742–56. Journal of Climate 16, 2232–2247.

Venema, V.K.C., Mestre, O., Aguilar, E., Auer, I., Guijarro, J.A., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., Zahradnicek, P., Viarre, J., Müller-Westermeier, G., Lakatos, M., Williams, C.N., Menne, M.J., Lindau, R., Rasol, D., Rustemeier, E., Kolokythas, K., Marinova, T., Andresen, L., Acquaotta, F., Fratianni, S., Cheval, S., Klancar, M., Brunetti, M., Gruber, C., Prohom Duran, M., Likso, T., Esteban, P., Brandsma, T., 2012. Benchmarking homogenization algorithms for monthly data. Climate of the Past 8, 89–115.

Vincent, L.A., Zhang, X., Bonsal, B.R., Hogg, W.D., 2002. Homogenization of daily temperatures over Canada. Journal of Climate 15, 1322–1334.

Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences, 2nd ed. Academic Press, Burlington, Massachusetts, USA 627 pp..

Williams, C.N., Vose, R.S., Easterling, D.R., Menne, M.J., 2006. United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data. ORNL/CDIAC-118, NDP-070. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee.