Surface Temperature Variations in East Africa and Possible Causes

JOHN R. CHRISTY, WILLIAM B. NORRIS, AND RICHARD T. MCNIDER

Earth System Science Center, University of Alabama in Huntsville, Huntsville, Alabama

ABSTRACT

Surface temperatures have been observed in East Africa for more than 100 yr, but heretofore have not been subject to a rigorous climate analysis. To pursue this goal monthly averages of maximum (T_{Max}), minimum (T_{Min}), and mean (T_{Mean}) temperatures were obtained for Kenya and Tanzania from several sources. After the data were organized into time series for specific sites (60 in Kenya and 58 in Tanzania), the series were adjusted for break points and merged into individual gridcell squares of 1.25°, 2.5°, and 5.0°.

Results for the most data-rich 5° cell, which includes Nairobi, Mount Kilimanjaro, and Mount Kenya, indicate that since 1905, and even recently, the trend of T_{Max} is not significantly different from zero. However, T_{Min} results suggest an accelerating temperature rise.

Uncertainty estimates indicate that the trend of the difference time series ($T_{Max} - T_{Min}$) is significantly less than zero for 1946–2004, the period with the highest density of observations. This trend difference continues in the most recent period (1979–2004), in contrast with findings in recent periods for global datasets, which generally have sparse coverage of East Africa.

The differences between T_{Max} and T_{Min} trends, especially recently, may reflect a response to complex changes in the boundary layer dynamics; T_{Max} represents the significantly greater daytime vertical connection to the deep atmosphere, whereas T_{Min} often represents only a shallow layer whose temperature is more dependent on the turbulent state than on the temperature aloft.

Because the turbulent state in the stable boundary layer is highly dependent on local land use and perhaps locally produced aerosols, the significant human development of the surface may be responsible for the rising T_{Min} while having little impact on T_{Max} in East Africa. This indicates that time series of T_{Max} and T_{Min} should become separate variables in the study of long-term changes.

1. Introduction

Because humanity lives on and obtains its sustenance from the surface of the earth, the near-surface air temperature is often viewed as a critical response variable associated with changes in forcing of the climate system. Several major efforts to create precise, long-term time series of near-surface air temperatures (or simply surface temperatures) have thus been carried out (Peterson and Vose 1997; Hansen et al. 1999; Brohan et al. 2006). However, problems are apparent in understanding the precision inherent in such compilations, especially when the time span for documenting changes increases to cover many decades.

Gridded global datasets of surface temperature may misrepresent trends in undersampled and poorly observed grid boxes (e.g., Christy et al. 2006). Undersampling occurs when observations are scarce or because much useful information has not yet been digitized. This is true of much of the African continent, and in particular of East Africa. Although a considerable amount of data are available for parts of the East African countries of Kenya and Tanzania, the available records are widely scattered, many records have not heretofore been converted to digital form, and the quality of the data is often poor, that is, either missing, illegible, or outside the range of possible values. These factors have prevented the full complement of available East African data from entering the databases of global datasets.

The goal of the research reported here is to contribute to a more thorough understanding of the twentieth-century climate of East Africa by constructing regional time series of temperatures from as many observations

Corresponding author address: John R. Christy, Earth System Science Center/Cramer Hall, University of Alabama in Huntsville, Huntsville, AL 35899.
E-mail: christy@nsstc.uah.edu

TABLE 1. Sources of data used in this study. Spans of years indicate only the first observation and the last observation of at least one station. Significant periods of observations are missing in most spans for many stations.

Monthly Kenya	No. of stations	T_{Max}	T_{Min}	T_{Mean}	Period
NCAR	25	X	X	X	1979–2004
British East Africa	77	X	X	X	1904–1974
GHCN	7	X	X	X	1911–2004
World Weather Records	15			X	1931–2000
Griffiths	3	X	X	X	1893–1962
NASA GISS	8			X	1895–2004
Monthly Tanzania					
NCAR	22	X	X	X	1979–2004
British East Africa	54	X	X	X	1922–74
GHCN	11	X	X	X	1875–2004
World Weather Records	9			X	1931–2000
Griffiths	2	X	X	X	1875–1962
NASA GISS	9			X	1892–2004

as possible for the period of 1905–2004. For climate studies this region is useful because at some locations observations began before 1900. The 5° grid cell with its southwest corner at 5°S, 30°E, and covering parts of Kenya and Tanzania, is of special interest. Not only does it contain major population centers of the two countries, but it also contains Mount Kilimanjaro and Mount Kenya. The shrinkage of the ice fields of these two mountains has been well documented (e.g., Thompson et al. 2002).

This paper begins by describing where the datasets were found and how they were organized to produce basic, unaltered time series (often more than one) for each station. Next, the paper describes how multiple series at a site were reduced to a single series, how inhomogeneities were detected and removed, how the resulting site-specific series were merged to create regional products, and how error estimates were created. The continued trend difference between maximum and minimum is contrasted with the few global datasets (e.g., Vose et al. 2005) that track minimum and maximum temperatures separately. Finally, hypotheses are offered that may explain notable features of the data. Following Pielke et al. (2007), we discuss the hypothesis that attempting to document changes in climate resulting from changes in forcings in the deep atmosphere (such as from enhanced greenhouse emissions) is better done by monitoring daily maximum temperatures than by daily means or minima.

2. Data sources

Table 1 lists the five sources of data that are used in this investigation, described below. The considered data quantities are the monthly means of the daily maximum and minimum temperatures (hereafter T_{Max} and T_{Min}, respectively). Two forms of averages of T_{Max} and T_{Min} were also considered, called T_{Avg} and T_{Mean}. The slight difference between them is discussed in section 5. After an extensive search for sources of data for Kenya and Tanzania, the following were found.

a. British East Africa summaries

From 1904 to 1974, the British East Africa (BEA) Meteorological Service, later the East African Meteorological Department, published annual summaries of monthly T_{Max} and T_{Min} for the temperature stations of Kenya. Beginning in the 1920s, after the Tanzanian mainland passed from German to British control, these summaries included Tanzania. Paper and/or electronic images of the annual summaries were obtained from the National Climatic Data Center (NCDC) and the U.K. Meteorological Library, and then were manually digitized at the University of Alabama in Huntsville (UAH). Annual summaries were available for most years since 1946, but many could not be located for the first half of the twentieth century. BEA data constitute the backbone of this investigation for years prior to 1974.

b. Global Historical Climate Network

NCDC has collected and digitized monthly average T_{Max} and T_{Min} values for more than 5000 temperatures stations around the world from a variety of sources for the Global Historical Climate Network (GHCN) project (Peterson and Vose 1997; Peterson et al. 1998). Time series of some of these stations were labeled "GHCN" in the archive. Others, identified as "Griffiths" (Table 1), were the files personally acquired through the efforts of Peterson and Griffiths (1996, 1997) in a notable data rescue effort focused on the earliest observations throughout Africa.

c. World Weather Records

Monthly values of T_{Mean} for Kenya and Tanzania were both published in the decadal World Weather Records (WWR) volumes and manually keyed at UAH, beginning with the 1931–40 edition (U.S. Department of Commerce 1949, 1959, 1967, 1985; Steurer 1993; Owen 1999). Data in the 1991–2000 edition included T_{Max} and T_{Min} values for a few stations, and these were also keyed for use. In the 1981–90 edition of WWR, the data for 1971–80 for three stations were republished along with their 1981–90 values. Over 50% of the republished values differed from the original values by at least 0.1°C. In the absence of knowing which version was better, both values were included among the samples to be examined by the best-guess algorithm (see section 4a).

d. National Center for Atmospheric Research

The National Center for Atmospheric Research (NCAR) maintains a digital archive of monthly summaries of worldwide stations beginning in 1979 based on information received electronically from synoptic reports. The ancillary data associated with the summaries includes the number of days per month on which observations were reported for each site. Both T_{Max} and T_{Min} values from this source were accepted when a minimum of 16 days was reported.

From January 1979 to June 1989 NCAR archived these data in one format, and from January 1987 to the present they were archived in another format. Comparison of the monthly values in the overlapping period revealed occasional slight differences, usually of 0.1°–0.2°C. When differences were found, the mean of the two representative values was calculated.

e. NASA Goddard Institute for Space Studies

The climate research group at the National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS) provides worldwide data of monthly T_{Mean} values (Hansen et al. 1999). The data used for this study were selected from the "raw GHCN data" files, but often differed from the data accessed directly from NCDC and labeled GHCN. The main differences between the two were monthly values listed as missing in one and available in the other. GISS often archived time series from more than one source (e.g., Tabora, Tanzania, was represented by five sources that appear to be multiple copies of WWR, but with varying missing months). Usually the source for GISS was either GHCN or WWR, but even then differences with the original sources appeared. Where differences from either GHCN or WWR were apparent, the GISS values were included as a separate source for T_{Mean} to be examined by the best-guess algorithm.

3. Data organization

The data from all sources were converted to a standard format and grouped by station name (the first significant organizational task). An attempt was then made to identify the particular instrument site associated with each time series from each source. In the case of the BEA series, monthly summary records were clearly site specific, with a given station name always referring to a unique location. In contrast, most non-BEA sources were multiple sites composited into long time series under a single name. Thus, the second significant organizational task was to examine each non-BEA time series and identify which portions were associated with a specific BEA instrument and/or site. This step led to a few series, initially associated with single station names, to be subdivided into as many as four segments, each corresponding to a unique BEA site.

A separate file was then created for each BEA site. For example, the BEA records for Mombasa, Kenya, were clearly identified as being from four locations through time. Thus, the Mombasa data were divided into four files with unique names, with each file listing all of the source data for that site. By organizing the data in this fashion, many of the potential biases introduced by station moves could be readily accounted for in our merging method (section 4) simply because the different thermometer shelter sites were treated as different stations even though the data as they were initially received were sometimes attributed to a single site.

The methodology described above, though thorough by many standards, very likely did not capture all station moves or other inhomogeneities and thus did not remove all discontinuities. Detecting these additional changes was not possible unless they were found in the break-point detection scheme (see section 4b).

Files resulting from the reorganization were not accepted unless at least 24 months of observations were available. Meeting the criteria were 60 separate collections of time series for Kenya and 58 for Tanzania.

4. Data-processing methods

Methods were applied to (a) reduce multiple time series for a single site to a single series, (b) identify and remove significant break points, and (c) merge time series for several sites in a region to a single time series representative of the region.

a. Best-guess algorithm

Using the data contributed by each source for a station (e.g., BEA, WWR, GHCN, etc.), a single "best-guess" value for each month was determined. Two simple statistics were computed to assist in this process. The first check was a simple calculation of the monthly anomaly relative to the monthly mean annual cycle for all time series at each station; call this mean annual cycle T_{Mean}. Also calculated was the standard deviation of the spread for each of the 12 months using all of the unadjusted times series for each station; call this T_{Mean}.

Data from the various sources tended to be in tight clusters. A cluster was defined as values within 0.03°C of each other. The average value of the cluster with the largest number of entries was calculated and tentatively accepted. In the case of competing clusters of the same number of entries, the cluster's average that was closest to the median of the mean annual cycle for the given

month was tentatively taken. While this has the tendency to reduce variability, the number of times this procedure was necessary was small because most stations were represented by only one or two time series, where in the latter case the two were essentially identical.

The departure of the tentatively selected value from T_{cycle} was then compared with the T_{SDcycle}, and if the resulting z score exceeded 1.96, then the value was discarded. While this procedure may eliminate a few true extremes, note that T_{SDcycle} was calculated from unadjusted data and thus represented a wider range than would be produced by the actual 95% confidence interval (CI) for a climate region with low natural interannual variability and accurate data. Very few values were eliminated this way. This ultimately resulted in a single T_{Max}, T_{Min}, and T_{Mean} series for each station.

b. Break-point detection

After the best-guess algorithm (BGA) generated a single time series for each of the 118 "stations," the data were examined for indications of break points, resulting from, for example, unrecorded station moves or other alterations. The goal was to discover the significant temperature shifts in each time series under the assumption that the remaining, undetected spurious shifts were essentially random in sign and small in magnitude. After the series were adjusted for the break points, it was assumed that each station's time series was an essentially "homogeneous segment" for merging into a regional time series (see Christy et al. 2006).

Break points were detected by applying a statistical test directly to each time series of anomaly observations. Using the method of Haimberger (2007), which is similar to that of Christy et al. (2006), the most consequential break points were identified.

The test statistic is

$$\tau_k = \Delta\{[\mu_k^-(\Delta) - \mu_k(2\Delta)]^2 \\ + [\mu_k^+(\Delta) - \mu_k(2\Delta)]^2\}/\sigma_k(2\Delta),$$

where k is the target month being examined for a break point, Δ is the length in months of a period before and a period after k, $\mu_k(2\Delta)$ is the mean, $\sigma_k(2\Delta)$ is the standard deviation of the target values over both periods, $\mu_k^-(\Delta)$ is the mean of the target values over the period before k, and $\mu_k^+(\Delta)$ is the mean of the target values over the period after k. Initially, Δ was chosen to be 36 months, but it was shortened to 24 months and then 12 months, as the endpoints of the series were approached. This process produced a time series of τ values. Break points were assumed to occur where maximum values of τ exceeded a specified threshold H. The break points

marked the points where the time series should be separated into "homogeneous segments."

The results of this paper are for time series parsed for thresholds of $H = \infty$ (i.e., no break points, and thus unadjusted data), 20, and 12. The value of 20 corresponds to a significance z score of about 4.5 (a temperature shift of about 1°C), and that of 12 corresponds to a z score of about 3 (a temperature shift of about 0.6°C). These three thresholds were used to help quantify the "parametric" uncertainty inherent in this dataset construction process.

Because the data availability for East Africa is inconsistent over the twentieth century, we report results of time series that begin in 1905, 1946, and 1979, with each ending in 2004. Data before 1905 turned out to be too sparse to be regionally meaningful. Data after 1945 represent the highest density of observations, and data after 1978 account for a period of significant population growth and urban migration in East Africa. Break-point detection was carried out separately for each starting date because the number and location of break points was a function of this date.

c. Merging methodology and gridcell averaging

Kenya and Tanzania can be essentially covered by a box having its southwest corner located at 10°S, 30°E, and its northeast corner at 5°N, 45°E. This box can be conveniently subdivided into nine nonoverlapping 5° cells. Each of these was subdivided into four 2.5° cells (36 in all), and each of these in turn was subdivided into four 1.25° cells (144 in all). These subdivisions provided the opportunity to examine regional trends at three levels of resolution—5°, 2.5°, and 1.25°.

After the individual station time series were adjusted according to the break-point methodology described above, they were merged into regional series, where each cell was taken to be a region. Each cell was associated with a circle of influence centered on the center point of the cell but that extended beyond the boundaries of the cell. A regional series was created by merging the series for stations within the circle, not just within the cell. Because the data are generally not dense in space and time, the larger circle allowed more stations to be combined and the noise to be reduced. Two radii were used for each gridcell size to help understand the parametric uncertainty in the dataset construction process. Specifically, the circle for the 1.25° cells was given radii of 100 and 200 km; for 2.5° cells, the radii were 200 and 300 km; and for 5° cells, the radii were 400 and 500 km.

Each cell was then checked to determine whether it contained at least one station. If not, no further processing of the cell was done, even if stations existed in

the cell's larger circle of influence. If the cell contained at least one station, then additional criteria were checked. Stations within the cell itself were tested to determine whether the valid temperatures collectively covered at least 75% of the months in the time period. If so, processing continued. If not, the stations were checked to determine whether the valid temperatures collectively covered 50% of the time period. If so, and if the valid temperatures of all of the stations in the circle of influence covered at least 90% of the time period, processing continued. If not, no further processing was done.

In cases when processing continued, the record of each station in the circle of influence was tested to make sure it overlapped the record of at least one other station so that debiasing could be carried out. If not, the station was omitted unless its record by itself accounted for at least 90% of the time period. After stations with unqualified records were omitted, the remaining stations were checked to make sure they still accounted for at least 90% of the time period. For cells having only one qualifying station, the trend for the cells was computed directly from the anomalies of the single time series. For cells having more than one station, the times series for these stations were merged into a single series. Merging followed the technique of Christy et al. (2006) in which times series are debiased relative to one another and combined to create a single regional series from which the trend was then computed.

The goal of varying the parameters was to create several regional time series for each cell to quantify variability or error. By varying the construction parameters and examining the spread of results, one measure of the confidence that may be ascribed to the results could be gained. For each time series the following methods were used for this purpose:

1) Where possible, a regional time series was created for each cell from the stations within the cell's circle of influence and the trend was computed. By calculating trends using two values for the radius of the circle and two values of H, four trends were produced. The method was applied to 1.25°, 2.5°, and 5° cells, though for the analysis to follow, only the four values from the full 5° grid box will be used.

2) For the central 5° cell two additional methods were used.

   (a) The median of the trends of the nested cells (either the 1.25° cells or the 2.5° cells) was computed. This method produced eight trends—four trends from the 1.25° cells (radius = 100, 200 km; H = 20, 12) and four from the 2.5° cells (radius = 200, 300 km; H = 20, 12).

   (b) After creating the time series of comparably nested cells the series were merged into a single series using the technique of Christy et al. (2006). Cells were comparable if they were the same size and were based on the same radius and the same value of H. The trend of the resulting anomaly series was computed. This method also produced eight trend values.

For the central 5° cell discussed below, we have a total of two, four, and four trend values, from methods 1, 2a, and 2b, respectively, for each H (either 12 or 20.) This gives a total of 10 realizations for each H and a total of 20 for both.

5. Trend results

Because the metric of the "linear trend of anomalies" is sensitive to homogenization methods, it exposes instances where the method has noticeable parametric dependence. The main results come from the central 5° cell.

The first example shows the influence of the choice of H. The three time series in Fig. 1 represent the T_{Max} anomalies since 1905 for the 2.5° cell whose southwest corner is 5°S, 35°E—the southwestern quarter of the central 5° cell. This cell uses data from six stations within the cell and eight outside, but within a 300-km radius of influence. In the merged, unadjusted time series (top, $H = \infty$) there are clear indications of spurious shifts, particularly in the early part of the record. When the individual stations are adjusted for $H = 20$ and merged, the major shifts are removed (middle). Removing the shifts in the stations that trip the $H = 12$ threshold produces a time series with a slightly positive trend (bottom). The early part of the record contains more variability, in part because of fewer stations reporting at that time.

Figures 2 and 3 show the trends in the 2.5° cells for T_{Max} and T_{Min} (radius of 300 km) as color tiles for the period of 1946–2004 for the three thresholds of H. The central 5° cell is outlined with a darker border. (Note: The three time series of Fig. 1 are represented in the lower-left corner of the central 5° cell in Fig. 2.) Stations utilized in the calculations are shown as filled circles, while stations that are not used are open. The cells on the periphery of the central 5° cell generally depend on only one or two stations, so that as the break points are applied, the trend may change dramatically. However, within the central cell, changes are less remarkable because of the presence of many more stations. The outer cells, for which fewer stations are available, reveal a wider range of variation in both figures.
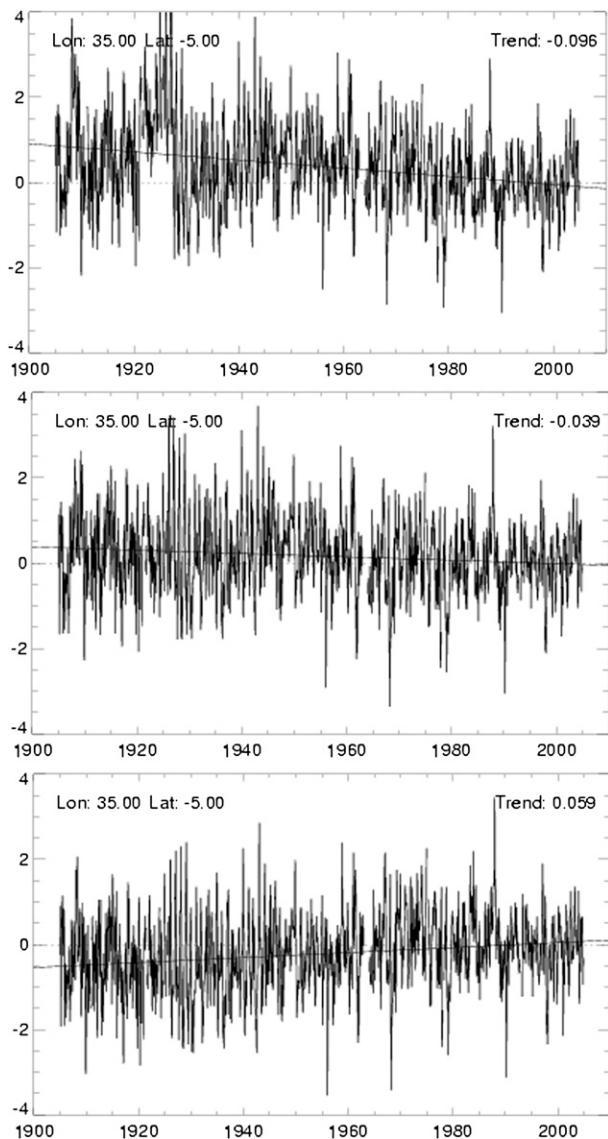
FIG. 1. T_{Max} time series of the 2.5° grid cell centered on 3.75°S, 36.25°E for three values of the break-point threshold: (top) $H = \infty$ (raw data), (middle) $H = 20$, and (bottom) $H = 12$.

Because stations tended to report T_{Min} more faithfully than T_{Max}, evidently because maximum thermometers failed more often as indicated by some reports, it was the case that more T_{Min} cells passed the criteria for calculation than T_{Max}. In general the results indicate the presence of near-neutral trends in T_{Max} and positive trends in T_{Min}.

Figure 4 displays the trend results for T_{Max} as a bar chart for the three time periods and the three threshold levels of H. For a given value of H, the error bars represent the ± 2-sigma extent determined from the 10 realizations based on parametric variations. In all cases,

the magnitude of the trend is close to zero, and only in the case of 1946–2004, unadjusted ($H = \infty$), does the error bar not cross zero. Because the pre-1946 data are sparse and variable, high confidence cannot be placed in the 1905–2004 trends.

Figure 5 displays the results for T_{Min}. As with T_{Max}, the 1905–2004 trends are not robust where large error bars indicate large parametric uncertainty. The trends in T_{Min} are very likely significantly positive and apparently accelerating as the magnitudes for the latest period exceed those of the 1946–2004 period. However, there is little parametric uncertainty in the remaining results. Adjustments for break points tend to reduce the trends, but they still remain positive in all of the later periods. Note that the magnitudes of the trends for $H = 20$ and $H = 12$ are quite similar in the 1946 and 1979 cases for both T_{Max} and T_{Min}, indicating the major breaks were captured with the weaker criterion ($H = 20$).

Finally, Fig. 6 shows the results of T_{Avg} and T_{Mean}. Here, T_{Avg} is generated only from stations with both T_{Max} and T_{Min} observations and is based primarily on BEA and GHCN data; T_{Mean} is generated from all data, including data from WWR, GISS, etc., for which only T_{Mean} are available. Thus, T_{Avg} is created from a subset of T_{Mean} listings and represents a time series in which the site location is known with greater confidence.

Included in the result are the trends calculated from Hadley Centre Climate Research Unit Temperature, version 3, variance adjusted (HadCRUT3v; land and ocean), University of East Anglia Climate Research Unit Temperature, version 3, variance adjusted (CRUTEM3v; land; Brohan et al. 2006), and GISS (Hansen et al. 1999). HadCRUT3v uses CRUTEM3v for the land surface stations, so the difference apparently is due to the contribution in HadCRUT3v from the sea surface temperatures in the southeast corner of the cell. (Trends calculated from HadCRUT3, i.e., the variance-unadjusted version, were essentially identical to HadCRUT3v, and so are not included.) The data in Fig. 6 show that the 1905–2004 trends of T_{Avg} and T_{Mean} are likely positive. The trends calculated in our study for 1946–2004 and 1979–2004 and across parameters are consistent with $+0.1$°C decade^{-1}. The surface temperature trends from HadCRUT3v and CRUTEM3v are similar to our analyses in the 1946–2004 period, but are much higher for 1979–2004, being highly inconsistent with the values we calculated under many parameterizations and data sources. GISS trends tend to be more positive than calculated by our methodology, especially in the most recent period as well, even though a type of urban adjustment has been applied (Hansen et al. 1999).

Only a few datasets have used separately minimum and maximum temperatures to track global trends in
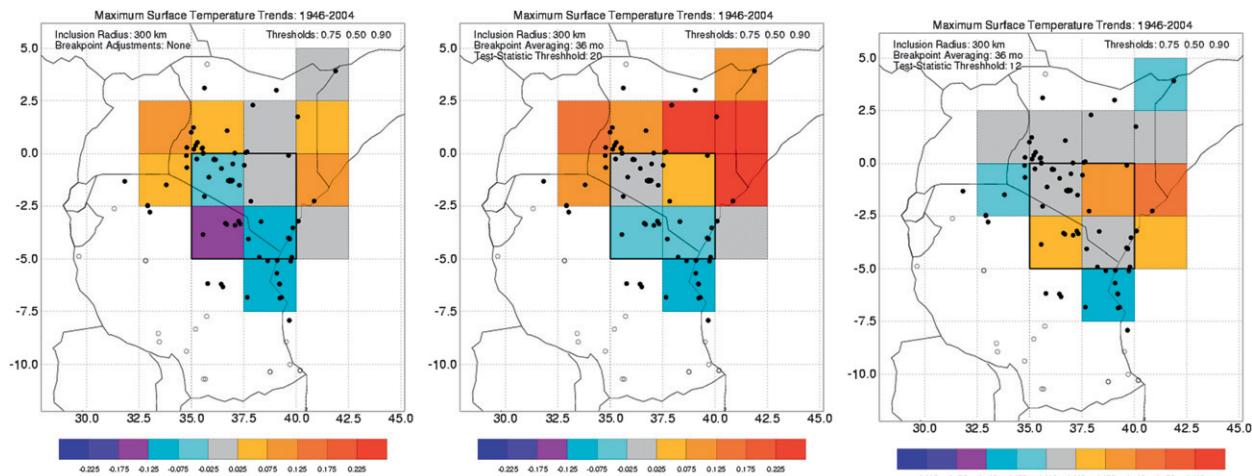
FIG. 2. Demonstration of the effect on 1946–2004 T_{Max} trends caused by varying the break-point detection threshold H applied to the station time series: (left) $H = \infty$ (raw data), (middle) $H = 20$, and (right) $H = 12$. The stations contributing to the time series of a 2.5° cell lie in a circle of radius 300 km centered on the cell. The 5° grid box outlined in bold is discussed in section 6.

temperature. The GHCN data or datasets do include maximum and minimum temperatures and have been important tools in understanding long-term climate trends. The GHCN data have shown large asymmetries in warming T_{Max} and T_{Min} through most of the data record, with warming in T_{Min} being more than twice that of T_{Max} (Karl et al. 1993). However, recent analyses of this dataset (Vose et al. 2005) have indicated that globally the asymmetry declined significantly in the 1979–2004 period, with T_{Max} and T_{Min} warming at nearly the same rate.

The large asymmetries in warming rates in T_{Max} and T_{Min} in the 1979–2004 in the present East Africa dataset are thus at odds with the global findings of Vose et al. (2005), and in fact the asymmetries are more like the global asymmetries found in the 1900–79 period re-

ported by Karl et al. (1993). There is a concern about the recent trends reported by Vose et al. (2005) in that the number of stations in the adjusted analysis drops from nearly 3000 stations in 1970 to less than 1500 in 2004. Further, the locations of stations are not well distributed around the globe, with few stations in the developing world in South America, India, and Africa, which are similar socioeconomically to East Africa.

6. Discussion

a. Trend error estimates

Though we have shown simple parametric error in the trends that were just calculated, assessing complete trend "error" is difficult. We first examine two
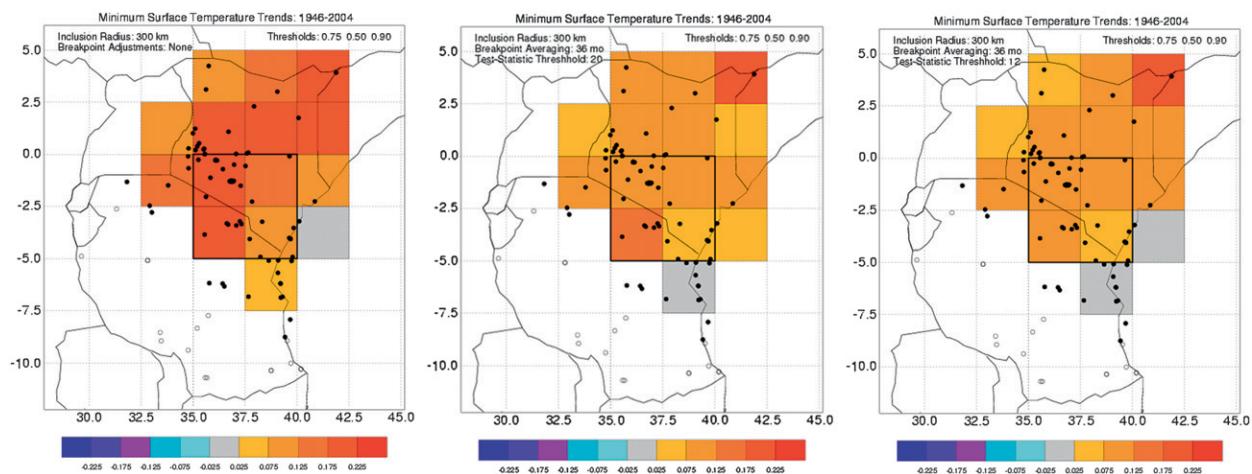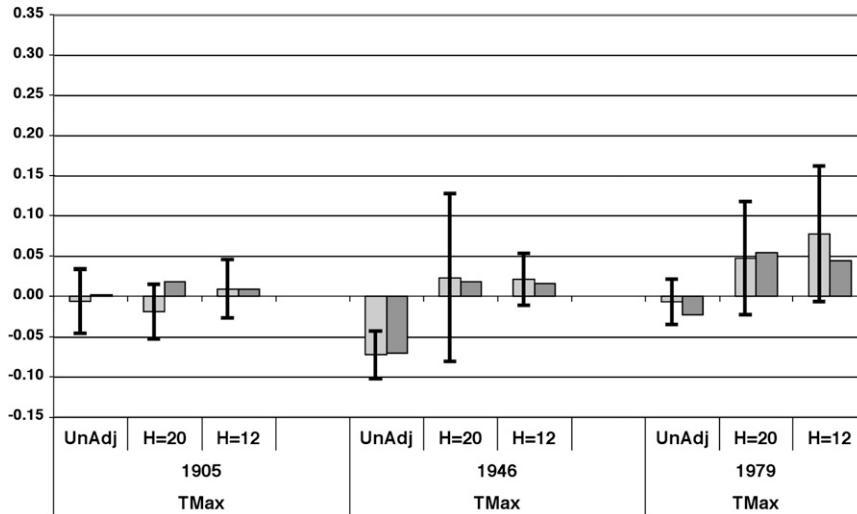

FIG. 3. Same as in Fig. 2, but for T_{Min}.

FIG. 4. Estimated value of the T_{Max} trend (°C decade^{-1}) beginning in the year indicated and ending in 2004 for the 5° grid box centered on 2.5°S, 37.5°E. Light gray represents the median of ten realizations based on parametric variations. The error bars are the $\pm 2\sigma$ spread of the 10 realizations. Dark gray represents the trend of the time series formed from stations within 400 km of the center of the cell. Break points for these stations were identified using a test statistic of $H = 20$.

calculations of "measurement error." This is error resulting from data problems, such as station moves, instrument changes, missing data, etc., and seeks to quantify a range of realizations that encompasses the true trend. The first measurement error model uses a traditional statistical approach, while the second looks at the structural uncertainty of the methodology, expanding on Figs. 4–6. We will also look at the issue of temporal sampling error, that is, the confidence one may place in the slope of a straight line fitted through variable data of a specific length of time.

For the first analysis, we calculate the year-by-year standard error of the individual time series values based on the multiple stations used to produce the mean values. Using the time series of the central 5° × 5° square, annual values of the standard error of T_{Mean}, T_{Max}, and T_{Min} are 0.28°, 0.41°, and 0.37°C, respectively, for the pre-1946 period and 0.18°, 0.24°, and 0.21°C, respectively, after 1945. Applying these calculated errors 1000 times by taking the original time series and perturbing each year by an error consistent with the distribution to generate 1000 time series, yields 95% confidence intervals for
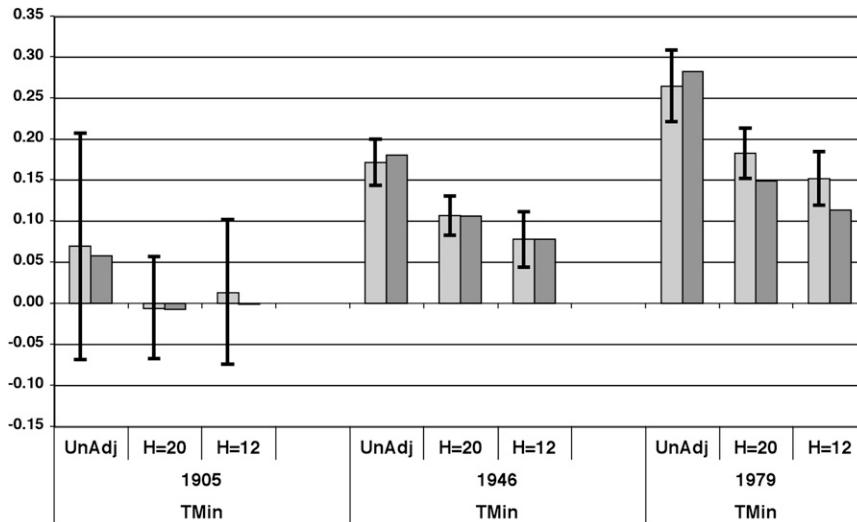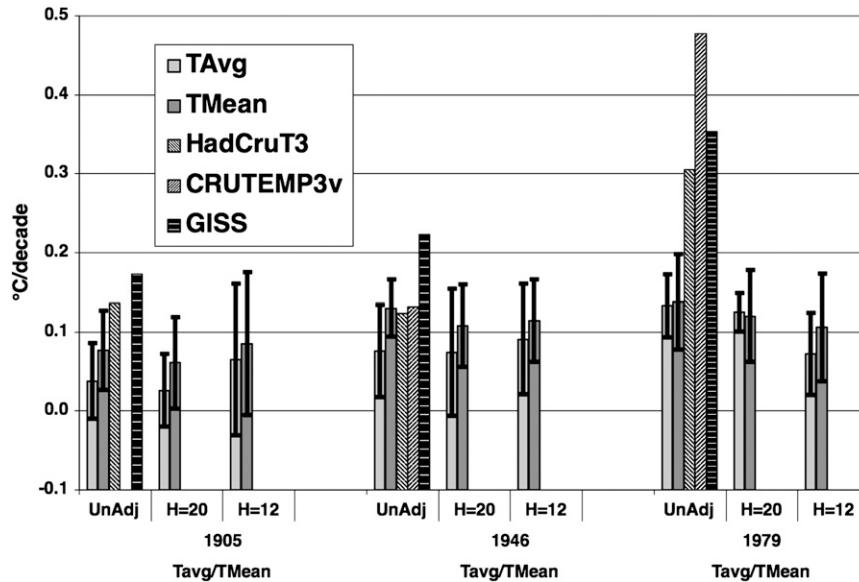

FIG. 5. Same as in Fig. 4, but for T_{Min}

Fig. 6. Same as in Fig. 4, but for T_{Avg} (light gray) and T_{Mean} (medium gray). The bars with patterns are trends calculated from the HadCRU3v, CRUTEM3v, and GISS surface temperature datasets.

trends (°C decade^{-1}), shown in Table 2 as the top error estimate. Because the errors are not large and are assumed to be random in time, they have little impact on the longest time periods.

The second error estimate in Table 2 targets "structural" uncertainty, quantified here by varying the parameters of the construction process (i.e., H and the radius of influence), as discussed previously. Because the errors calculated from 1905 show greater and, in our view, more realistic magnitudes with this second method, we believe it more accurately represents the measurement uncertainty of the trends. However, in Table 2, we take the parametric error range to be the 95% CI of the *combined* realizations for $H = 20$ and $H = 12$, and thus generally larger than that shown in Figs. 4–6 where values for $H = 20$ and $H = 12$ are separately calculated and presented. These estimates are listed as the middle error in Table 2. The general result regarding error distributions to this point is that as the time period shortens and the statistical error range increases, while the parametric error range may not. This occurs in the latter case because as the density of observations increases (toward the end), the influence of parametric variations on a larger sample introduces reduced changes in the mean of the larger sample.

A different type of error is now discussed. The trend used here is the slope of the line calculated by linear regression through the time series. If the time series is short, or if the individual values have large variations relative to the magnitude of the slope, then adding a

single anomalous value at the beginning or end of the data has the potential to tilt the line of best fit from its initial value and thus expose the time series as not robust to temporal sampling. This is not measurement error because a perfect time series will exhibit this characteristic, and will be called "temporal sampling error" here. In our case, temporal sampling error will have its greatest impact on the shortest period studied, the 26 yr during 1979–2004, because the longer periods will produce only minor temporal sampling error potential (Table 2, bottom error estimate).

The difficulty in calculating the temporal sampling error for a relatively short time series is that the theoretical population of 26-yr samples from which we theoretically extract samples at random must have experienced the same sequence of external forcing events as the observations (e.g., El Niño, volcanoes, solar, greenhouse gases, boundary layer influences, etc.). Not having access to such a population, we show the values in Table 2 assuming random forcing, which produces very large error ranges, likely much larger than those in reality.

b. Difference between T_{Max} and T_{Min} trends

The time series of annual anomalies of T_{Max} and T_{Min} for the central 5° cell are shown in Fig. 7 using the parameters of the "best estimate," that is, the trend that is calculated using the 5° grid as a single cell with $H = 20$ and the radius of influence of 400 km. The clear rise in T_{Min} is evident since 1946, whereas little change is evident in the time series of T_{Max} over the same period.

TABLE 2. Estimated trends (by least squares linear regression) and errors of the central 5° cell (°C decade^{-1}). Upper trend value: median of trends and 95% CI from 20 realizations of time series using both break-point thresholds ($H = 12$ and 20) and all of the subsetting variations described in section 6a. Lower trend value (italics): the "best estimate" trend determined by considering the 5° cell as a single grid square with a break-point threshold of $H = 20$ and a radius of 400 km. The three error estimates for each trend value are described in section 6a: (top) standard measurement error, (middle) parametric measurement error, and (bottom) temporal sampling error. The last rows show the T_{Mean} trends for this cell derived from the surface temperature datasets.

	1905–2004		1946–2004		1979–2004	
	Trend	Error	Trend	Error	Trend	Error
T_{Max}	+0.00	±0.01	+0.02,	±0.02	+0.05,	±0.09
	+0.02	±0.05	*+0.02*	±0.08	*+0.05*	±0.08
		±0.02		±0.05		±0.19
T_{Min}	+0.00	±0.02	+0.09	±0.03	+0.16	±0.11
	−0.01	±0.08	*+0.11*	±0.04	*+0.15*	±0.08
		±0.04		±0.05		±0.17
$T_{Max} - T_{Min}$	+0.00,	±0.02	−0.06	±0.03	−0.10	±0.08
	+0.03	±0.07	*−0.09*	±0.09	*−0.09*	±0.08
		±0.03		±0.05		±0.20
T_{Avg}	+0.04,	±0.01	+0.09	±0.02	+0.10	±0.06
	+0.02	±0.08	*+0.07*	±0.08	*+0.12*	±0.08
		±0.03		±0.04		±0.18
T_{Mean}	+0.06	±0.01	+0.11	±0.02	+0.12	±0.06
	+0.04	±0.08	*+0.10*	±0.06	*+0.11*	±0.07
		±0.03		±0.05		±0.18
HadCRUT3v	+0.14		+0.11		+0.31	
CRUTEM3v	-		+0.13		+0.47	
GISS	+0.17		+0.22		+0.35	

Evidence in Table 2 indicates that trends in T_{Mean}, T_{Min}, and ($T_{Max} - T_{Min}$) are very likely significantly different from zero for 1946–2004 and likely are so for 1979–2004. The magnitude of the T_{Max} trend is likely not significantly different from zero. The large asymmetries in warming rates in T_{Max} and T_{Min} since 1979 in the present analysis are thus at odds with Vose et al. (2005), who found comparable warming rates in T_{Max} and T_{Min} at the global scale for that period. Notably, however, the Vose et al. (2005) analysis was hampered by a lack of data for developing world in general (including East Africa).

The recent trends of T_{Mean} calculated from global datasets do not agree with our results for this cell. As shown in Table 2, the 1979–2004 T_{Mean} trend of the central 5° cell as produced by HadCRUT3v, CRUTEM3v, and GISS (+0.31°, +0.47°, and +0.35°C decade^{-1}, respectively) are markedly inconsistent with all of the time series for that cell constructed in this study. Evidently, the main signal used by HadCRUT3v for this cell since 1979 is derived from the single Nairobi, Kenya, station at Jomo Kenyatta Airport (P. Jones 2004, personal communication). Our unadjusted time series for this site does indeed show significant warming since 1979 (+0.25°C decade^{-1}), but the higher trend is not corroborated by the many nearby stations used in our analysis. Such differences were also found in central California (Christy et al. 2006) and northern Alabama (Christy 2002), where our more comprehensive reconstructions were on average about 0.1°C decade^{-1} more negative in the cells covering those areas versus values for the cell from global databases.

c. Possible causes for T_{Max} and T_{Min} differences

The fact that the trends in the two temperature measurements (T_{Max} and T_{Min}) are likely significantly different encourages an examination of the causes for the warming of T_{Min} and the significance of trends in T_{Min} in the context of tracking global climate change. Given a lack of detail on station siting and uncertainties in specifics on the boundary layer in East Africa, definitive reasons for the trends may not be available. However, general aspects of boundary layer behavior may provide some guide for interpreting the trends. Thus, the following should be viewed as a context and hypothesis for the trend differences that deserve discussion and further attention.

1) GENERAL DISCUSSION

What do differing T_{Max} and T_{Min} trends mean in the context of detecting the magnitude of anthropogenic warming? Usually T_{Max} occurs in the daytime when the surface is vertically connected via dry-adiabatic mixing processes to a mixing depth of 1.5–2.5 km as determined from the Nairobi radiosonde. The southeast corner of the central 5° cell is exposed to the trade wind inversion off the Indian Ocean and may not achieve such mixing depths. However, most of the cell, and indeed the area where most observations were made, is in the central highlands with relatively deep daytime mixing. Because of the vigorous mixing processes in the daytime, vertical gradients in potential temperatures are modest. Thus, T_{Max} is more representative of temperatures aloft, at least at the top of the boundary layer, and is better able to serve as a proxy for heat content of a substantial mass of the lower atmosphere.

On the other hand, T_{Min}, which occurs during the night or early morning, represents the temperature of a much smaller mass of air because nocturnal boundary layers (NBL) are often only a few hundred meters thick (Stull 1988). Because of smaller turbulent intensities substantial vertical gradients exist. Thus, T_{Min} is often representative only of a shallow layer and is sensitive to measurement heights and local land surface properties because of the strong vertical gradients (Runnalls and Oke 2006; Pielke and Matsui 2005). Also, wind speeds
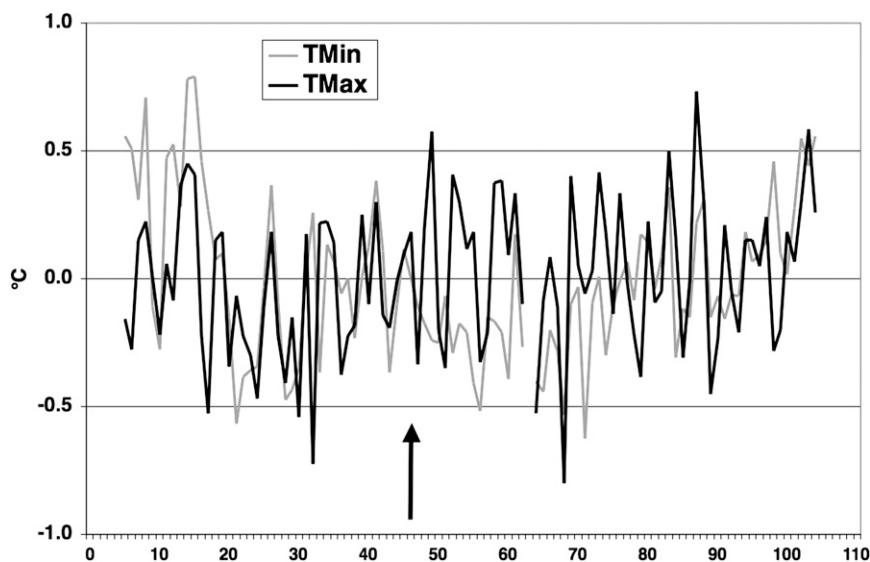
Fig. 7. Time series of annual anomalies of T_{Max} and T_{Min} as determined for the 5° Nairobi grid cell. The time series were formed from stations within 400 km of the center of the cell. Break points for these stations were identified using a test statistic of $H = 20$. The arrow indicates 1946, the year when a significant amount of data began to be available. Earlier anomalies, especially before 1922, are highly uncertain.

in the NBL are usually less than those in the daytime so that the horizontal footprint of the observation of T_{Min} is less than that of T_{Max}. Overall, the representative observational volume for detecting accumulated heat in the atmosphere is much smaller for T_{Min} than for T_{Max} (Pielke et al. 2007).

To further complicate matters for T_{Min}, McNider et al. (1995), Van de Wiel et al. (2002), and Walters et al. (2007) demonstrate that the NBL in which T_{Min} is measured acts as a delicate, nonlinear dynamical system. In some parameter spaces this system responds with large changes in T_{Min} for only slight changes in parameters, such as roughness, wind speed, or radiative forcing. Colder T_{Min} temperatures occur when the stable boundary layer decouples from the deep layer above and cools radiatively. Warmer T_{Min} temperatures are maintained when the surface is coupled by turbulent mixing to the warmer layer aloft. Any slight forcing that disrupts the decoupling leads to greater mixing and thus to warmer values of T_{Min}. Thus, T_{Min} is often more dependent on the turbulent state of the atmosphere than on temperature imposed from the atmosphere above (McNider et al. 1995).

As shown by Shi et al. (2005) and Runnalls and Oke (2006), there are many candidates for increasing the frequency of disruption events of the stable NBL, including changes in roughness (with the introduction of buildings or trees), surface thermal forcing (with the introduction of heat absorbing surfaces such as asphalt),

heat capacity of the surface (with irrigated cropland replacing desert soil), thermal forcing from aerosols in the shallow layer, and greenhouse gas increases (Walters et al. 2007). Such disruptive events need to occur only a few times more per year than previously observed to produce a noticeable change in the average T_{Min} because these transition events often warm the air by several degrees. This warming is caused by a transition to a more turbulent state in which heat is redistributed. Even if the transition to a more turbulent state is by greenhouse gas radiation (Walters et al. 2007), the resulting warming is due to a redistribution of heat to the lowest few hundred meters of the atmosphere rather than reflecting an accumulation of heat in the deep atmosphere.

As noted in Pielke et al. (2007) and Lin et al. (2007), because T_{Min} represents only a shallow layer and trends in T_{Min} could be more indicative of trends in the local turbulent state, then a better proxy for detecting climate change of the deep, global atmosphere may be found in T_{Max}. Though T_{Max} is also influenced by variations in local surroundings (e.g., by cooling from irrigation or warming from urbanization), the greater ventilation and connectivity to the deeper troposphere argues that it is a better proxy than T_{Mean} (influenced by T_{Min} for detecting changes considered to be affecting the deep troposphere, such as enhanced greenhouse warming. This is not meant to say that T_{Max} will always be a good measure for deep atmosphere trends, because upper-level inversions can also cause the daytime boundary

layer temperatures to be disconnected from the boundary layer, but it has a higher probability of being representative of the deeper atmosphere than T_{Min}. Confounding factors include alterations of the surface through urbanization, agriculture, etc.

2) APPLICATION TO THE CENTRAL 5° CELL

The main grid cell in this study contains the highlands of Kenya and Tanzania, and in particular those towns and cities with weather stations where tremendous growth and change have occurred. Nairobi, for example, consisting of 714 km^2, has experienced significant surface changes resulting from development. Urban areas expanded from 14 to 61 km^2 and agricultural lands expanded from 50 to 88 km^2 between 1976 and 2000. Forested areas declined from 100 to 24 km^2 (Mundia and Aniya 2005). This growth has altered the landscape and some of its meteorological parameters such as roughness and heat capacity that may retard or prevent the decoupling of the NBL, and thus lead to warmer surface conditions at night.

For example, changes in roughness can dramatically change surface temperatures in the stable boundary layer. As shown by McNider et al. (1995), as trees or buildings replace grass, increases in roughness can lead to substantially warmer temperatures. Also, for low and moderate wind speeds, increases in heat capacity arising from concrete replacing vegetative mulch or irrigation increasing soil water content (with accompanying increases in heat capacity and conductivity), can lead to perceived warming (Shi et al. 2005). Given the rapid population growth near the observation sites in the East African highlands, such changes are likely.

In addition to land use change, aerosol forcing in the NBL may play a role. In East Africa the common practice of burning biomass for warmth, cooking, and light, especially in the early evening, tends to fill the shallow NBL of these communities, where most weather stations are sited, with a visible layer of smoke. Additionally, the large smoke aerosols, smaller hygroscopic aerosols, and larger coated organic aerosols may readily swell when the humidity reaches 80% (common in East African evenings). In combination, these produce a nighttime pall that is characteristic of the underdeveloped world. Although the magnitude of aerosol forcing in East Africa is not known, model studies in Los Angeles, California, where the concentration of large thermally active aerosols is likely much smaller than East Africa, showed that the presence of aerosols accounted for an enhancement of nocturnal downwelling radiation of 13 W m^{-2} (Jacobson 1997). A recent study in India, where aerosol forcing may be similar to East Africa, attempted to account for their role and estimated a daily mean of downwelling radiation enhancement from aerosols of 6.5 to 8.2 W m^{-2} (Panicker et al. 2008.)

As shown by Eastman et al. (2001), such additional forcing from greenhouse gases may differentially act to warm the decoupled NBL because the additional forcing is confined to a shallow layer. In a sensitivity study of direct temperature response using techniques of nonlinear analysis to greenhouse gas forcing, Walters et al. (2007) showed that the NBL had a range of sensitivity depending on the imposed parameter space.

Figure 8 shows a bifurcation diagram for enhanced downward radiation from aerosols, which is like the downward radiation from greenhouse gases (see Nair et al. 2008, manuscript submitted to *J. Geophys. Res.*, hereafter NAI). It shows (a) large linear temperature sensitivity in light winds, (b) multiple solutions for intermediate winds, and (c) less sensitivity in strong winds. Under light winds (Fig. 8a) as downward radiation increases the NBL temperature increases linearly in response with a slope or sensitivity of about 0.12 K (W m^{-2})$^{-1}$. Under strong winds (Fig. 8c), when the NBL depth is greater and mixing is strong, the simple model indicates less sensitivity because temperatures stay warm due to mixing. However, at intermediate winds (Fig. 8b), the temperature difference between the two states can be of the order of 7–9 K and have a sensitivity of 0.28–0.36 K (W m^{-2})$^{-1}$, depending on roughness length. This large sensitivity is due to a dynamic feedback in that the additional downward radiative forcing destabilizes the NBL and disrupts the decoupling process, leading to large increases in T_{Min} as warm air is mixed from aloft. Note that this warming is due to a redistribution of heat, *not* to heat added by the downward radiation. While the two-layer boundary layer model used to develop Fig. 8 is simple and the sensitivity depends partially on the assumed layer depth, the layer depth is chosen such that the model replicates the actual difference in surface air temperatures in observations between a windy (weakly stable) and calm (strongly stable) night (Steeneveld et al. 2006). While the magnitude of the warming resulting from destabilization may be different between the real atmosphere and the simple model, the process of radiative destabilization by weak radiative forcing is likely real.

Recent results to be reported (NAI) using a more complete model show a sensitivity for the NBL of 0.06 K (W m^{-2})$^{-1}$ for a light wind case in general agreement with the simple model. NAI also show that the sensitivity of the daytime (convective boundary layer) temperature to aerosol forcing was only one-third as sensitive as the NBL. Thus, while the change in incoming shortwave energy resulting from aerosols was larger than the enhanced downward longwave energy
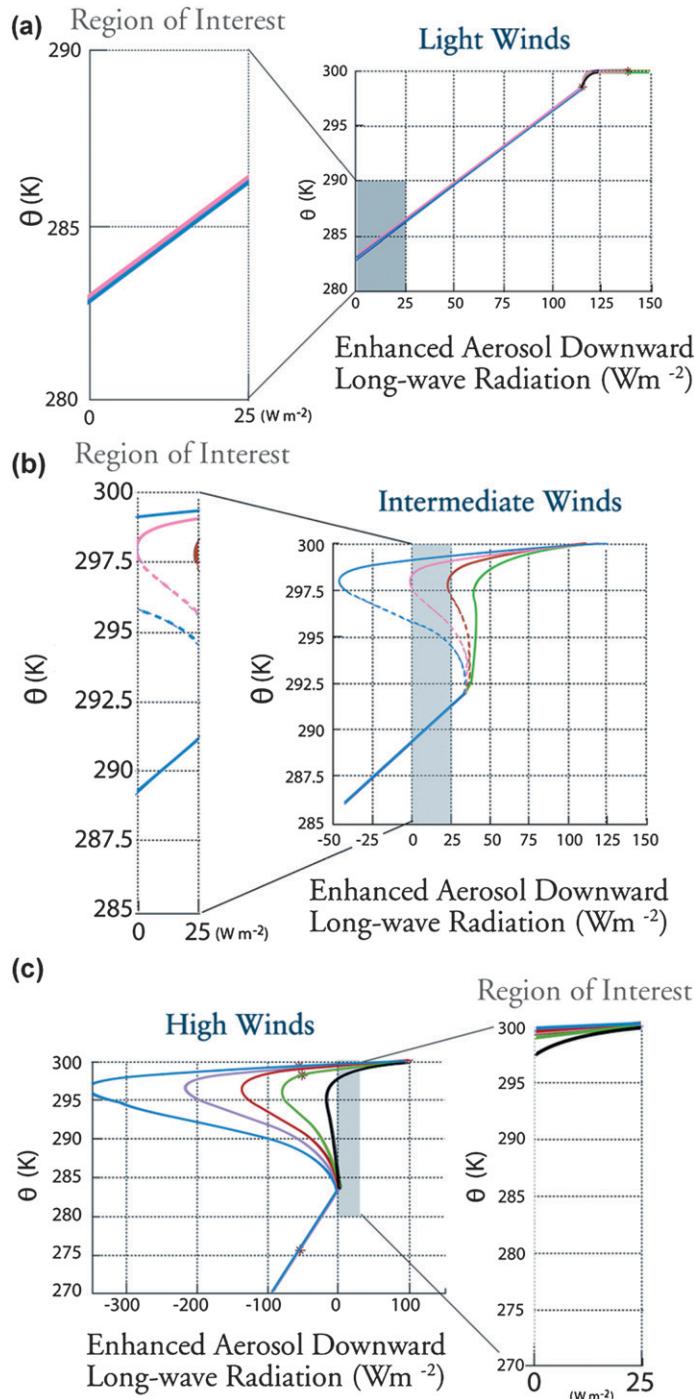
FIG. 8. Bifurcation diagrams of the temperature solutions for variations in radiative forcing (bifurcation parameter) of the nocturnal boundary layer or NBL (Walters et al. 2007). The regions of interest are the 0–25 W m^{-2} anomalous (aerosol) forcing calculations, which are enlarged. (a) Light wind case (geostrophic speed 3 m s^{-1}) showing no impact on the turbulent state, with NBL temperatures warming linearly with forcing. (b) Intermediate wind case (7 m s^{-1}) in which solutions can be much warmer when the decoupling process is interrupted and mixing occurs. (c) High-wind case (10 m s^{-1}) when mixing is strong generating consistently warm NBL temperatures. Line colors give roughness length: $z_0 = 0.1$ m (green), $z_0 = 0.25$ m (red), $z_0 = 0.5$ m (pink), and $z_0 = 1.0$ m (blue).

from aerosols, the net effect was little difference in the T_{Min}. While Walters et al. (2007) showed that direct greenhouse gas forcing or clouds can cause similar effects, the magnitude of the aerosol forcing in the heavily polluted environments likely cannot be ignored.

Increases in cloudiness can also impact nighttime temperatures (Dai et al. 2006). However, inferred trends in East Africa cloudiness in studies of glacial loss on Mount Kilimanjaro indicate a possible decrease (Mote and Kaser 2007). In summary, it seems probable that the T_{Min} signal observed in East Africa is due to, at least in part, the changing surface character and/or air quality in the NBL and its influence on mixing from warmer layers above.

The observation that the trends of T_{Max} and T_{Min} are different in many locations around the world has been documented in several sources (e.g., Easterling et al. 1997) though when globally averaged, the difference seems to have been decreasing in the most recent two decades (Vose et al. 2005), unlike our results for East Africa. The implication of our reconstruction for East Africa, north Alabama (Christy 2002), and central California (Christy et al. 2006) support the lack of positive trends in T_{Max}, and thus the possibility that these truly indicate the nature of changes in the deep troposphere. The reason it is important to separate the shallow warming as measured by T_{Min} from warming of the deep atmosphere is that the heat accumulated in the deep atmosphere is responsible for many of the indirect climate feedbacks, such as that of increased water vapor. The increased radiative path-length for water vapor resulting from warming of a shallow layer is not likely to support the level of feedback in a deep, warmer atmosphere.

7. Conclusions

Constructing a dataset of surface temperatures for East Africa requires significant human intervention to digitize and to make decisions about such basic activities as organizing the data into site-consistent time series. Once the data were organized by site (60 in Kenya and 58 in Tanzania), it was found that for many sites, multiple sources of data existed. A "best-guess algorithm" was applied to the sources to achieve a single time series for each of the 118 sites. A statistical method was then applied to detect and remove inhomogeneities at differing thresholds. From these 118 time series regional time series were created that combined individual time series meeting criteria of nearness to the region, length of record, and the ability to be merged with the other time series.

The most data-rich region studied here was the 5° grid cell bounded by 5°S–0° and 35°E–40°E, which represents portions of southern Kenya and northeastern Tanzania. For the 100-yr period from 1905 to 2004 in this grid cell, the trends were near zero for both T_{Max} and T_{Min}, but confidence in these results is low because of the relatively sparse data in the years before 1946. Beginning with 1946 and ending in 2004, near-zero trends were found for T_{Max}. The T_{Min} trends were more positive, and significantly so based on both measurement error and temporal sampling error. It is difficult to assess the measurement error of these trends, but using the spread of 20 realizations in which the construction parameters were varied, the range of $\pm0.10°C$ decade^{-1} is plausible.

The fact that the difference in trends in T_{Max} and T_{Min} continues, and in fact accelerates, in the period of 1979–2004 in East Africa may be important in interpreting the results of Vose et al. (2005). While it is possible that East Africa difference trends are indeed different than that of the globe as provided by Vose et al. (2005), there is concern that the reduced number of stations in the 1979–2004 GHCN dataset may not be sampling many of the areas of the globe that are behaving like East Africa. Thus, it is important that the GHCN dataset be expanded to include more stations distributed around the globe.

The noticeable difference in trends of T_{Max} and T_{Min} implies that daytime and nighttime temperatures are responding differently to environmental factors. Changes in the surface characteristics and the boundary layer atmospheric constituents may be responsible for the relatively recent and rapid rise in T_{Min}. There appears to be little change in East Africa's T_{Max}, and if T_{Max} is a suitable proxy for climate changes affecting the deep atmosphere, there has been little impact in the past half-century.

The investigation of the surface temperature record as an indicator of human-induced climate change involves understanding the complex behavior of boundary layer processes (where surface temperatures are actually measured) and how temperatures within it are affected by the numerous changes that occur. This is an area of research open for considerable inquiry because it raises new questions concerning the types of data indices now used to detect climate change. At the least, the time series of both T_{Max} and T_{Min} should become separate variables to be studied for long-term changes.

REFERENCES

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperatures changes: A new data set from 1850. *J. Geophys. Res.,* **111,** D12106, doi:10.1029/2005JD006548.

Christy, J. R., 2002: When was the hottest summer? A State Climatologist struggles for an answer. *Bull. Amer. Meteor. Soc.,* **83,** 723–734.

——, W. B. Norris, K. Redmond, and K. P. Gallo, 2006: Methodology and results of calculating central California surface temperature trends: Evidence of human-induced climate change? *J. Climate,* **19,** 548–563.

Dai, A., T. R. Karl, B. Sun, and K. E. Trenberth, 2006: Recent trends in cloudiness over the United States: A tale of monitoring inadequacies. *Bull. Amer. Meteor. Soc.,* **87,** 597–606.

Easterling, D. R., and Coauthors, 1997: Maximum and minimum temperature trends for the globe. *Science,* **277,** 364–367.

Eastman, J. L., M. B. Coughenour, and R. A. Pielke, 2001: The effects of CO_2 and landscape change using a coupled plant and meteorological model. *Global Change Biol.,* **7,** 797–815.

Haimberger, L., 2007: Homogenization of radiosonde temperature time series using innovation statistics. *J. Climate,* **20,** 1377–1403.

Hansen, J., R. Ruedy, J. Glascoe, and M. Sato, 1999: GISS analysis of surface temperature change. *J. Geophys. Res.,* **104,** 30 997–31 022.

Jacobson, M. Z., 1997: Development and application of a new air pollution modeling system—Part III. Aerosol-phase simulations. *Atmos. Environ.,* **31,** 587–608.

Karl, T. R., and Coauthors, 1993: Asymmetric trends in surface temperature. *Bull. Amer. Meteor. Soc.,* **74,** 1007–1023.

Lin, X., R. A. Pielke Sr., K. G. Hubbard, K. C. Crawford, M. A. Shafer, and T. Matsui, 2007: An examination of 1997–2007 surface layer temperature trends at two heights in Oklahoma. *Geophys. Res. Lett.,* **34,** L24705, doi:10.1029/2007GL031652.

McNider, R. T., X. Shi, M. Friedman, and D. E. England, 1995: On the predictability of the stable atmospheric boundary layer. *J. Atmos. Sci.,* **52,** 1602–1614.

Mote, P. W., and G. Kaser, 2007: The shrinking glaciers of Kilimanjaro: Can global warming be blamed? *Amer. Sci.,* **95,** 318–325.

Mundia, C. N., and M. Aniya, 2005: Analysis of land use/cover changes and urban expansion of Nairobi city using remote sensing and GIS. *Int. J. Remote Sens.,* **26,** 2831–2849.

Owen, T. W., Ed., 1999: *Africa.* Vol. 5, *World Weather Records 1981–90,* U.S. Department of Commerce, NOAA/NESDIS, 343 pp.

Panicker, A. S., G. Pandithurai, P. D. Safai, and S. Kewat, 2008: Observations of enhanced aerosol longwave radiative forcing over an urban environment. *Geophys. Res. Lett.,* **35,** L04817, doi:10.1029/2007GL032879.

Peterson, T. C., and J. F. Griffiths, 1996: Colonial era archive data project. *Earth Syst. Monit.,* **6,** 8–16.

——, and ——, 1997: Historical African data. *Bull. Amer. Meteor. Soc.,* **78,** 2869–2872.

——, and R. S. Vose, 1997: An overview of the Global Historical Climatology Network temperature database. *Bull. Amer. Meteor. Soc.,* **78,** 2837–2849.

——, ——, R. Schmoyer, and V. Razuvaëv, 1998: Global Historical Climatology Network (GHCN) quality control of monthly temperature data. *Int. J. Climatol.,* **18,** 1169–1179.

Pielke, R. A., Sr., and T. Matsui, 2005: Should light wind and windy nights have the same temperature trends at individual levels even if the boundary layer averaged heat content change is the same? *Geophys. Res. Lett.,* **32,** L21813, doi:10.1029/2005GL024407.

——, and Coauthors, 2007: Unresolved issues with the assessment of multidecadal global land temperature trends. *J. Geophys. Res.,* **112,** D24S08, doi:10.1029/2006JD008229.

Runnalls, K. E., and T. R. Oke, 2006: A microclimatic technique to detect inhomogeneities in historical records of screen-level air temperature. *J. Climate,* **19,** 959–978.

Shi, X., R. T. McNider, D. E. England, M. J. Friedman, W. Lapenta, and W. B. Norris, 2005: On the behavior of the stable boundary layer and role of initial conditions. *Pure Appl. Geophys.,* **162,** 1811–1829.

Steeneveld, G. J., B. J. H. van de Wiel, and A. A. M. Holtslag, 2006: Modeling the evolution of the atmospheric boundary layer coupled to the land surface for three contrasting nights in CASES-99. *J. Atmos. Sci.,* **63,** 920–935.

Steurer, P. M., Ed., 1993: *Africa.* Vol. 5, *World Weather Records 1971–80,* U.S. Department of Commerce, NOAA/NESDIS, 465 pp.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology.* Kluwer Academic, 670 pp.

Thompson, L. G., and Coauthors, 2002: Kilimanjaro ice core records: Evidence of Holocene climate change in tropical Africa. *Science,* **298,** 589–593.

U.S. Department of Commerce, 1949: *World Weather Records 1931–40.* U.S. Weather Bureau, U.S. Government Printing Office, 646 pp.

——, 1959: *World Weather Records 1941–50.* U.S. Weather Bureau, U.S. Government Printing Office, 1361 pp.

——, 1967: *Africa.* Vol. 5, *World Weather Records 1951–60,* ESSA, U.S. Government Printing Office, 545 pp.

——, 1985: *Africa.* Vol. 5, *World Weather Records 1961–70,* NOAA/NESDIS, 533 pp.

Van de Wiel, B. J. H., A. F. Moene, R. J. Ronda, H. A. R. De Bruin, and A. A. M. Holtslag, 2002: Intermittent turbulence and oscillations in the stable boundary layer over land. Part II: A system dynamics approach. *J. Atmos. Sci.,* **59,** 2567–2581.

Vose, R. S., D. R. Easterling, and B. Gleason, 2005: Maximum and minimum temperature trends for the globe: An update through 2004. *Geophys. Res. Lett.,* **32,** L23822, doi:10.1029/2005GL024379.

Walters, J. T., R. T. McNider, X. Shi, W. B. Norris, and J. R. Christy, 2007: Positive surface temperature feedback in the stable nocturnal boundary layer. *Geophys. Res. Lett.,* **34,** L12709, doi:10.1029/2007/GL029505.