A Technique to Detect Microclimatic Inhomogeneities in Historical Records of Screen-Level Air Temperature

K. E. Runnalls* and T. R. Oke

Department of Geography, University of British Columbia, Vancouver, British Columbia, Canada

(Manuscript received 17 December 2004, in final form 6 June 2005)

ABSTRACT

A new method to detect errors or biases in screen-level air temperature records at standard climate stations is developed and applied. It differs from other methods by being able to detect microclimatic inhomogeneities in time series. Such effects, often quite subtle, are due to alterations in the immediate environment of the station such as changes of vegetation, development (buildings, paving), irrigation, cropping, and even in the maintenance of the site and its instruments. In essence, the technique recognizes two facts: differences of thermal microclimate are enhanced at night, and taking the ratio of the nocturnal cooling at a pair of neighboring stations nullifies thermal changes that occur at larger-than-microclimatic scales. Such ratios are shown to be relatively insensitive to weather conditions. After transforming the time series using Hurst rescaling, which identifies long-term persistence in geophysical phenomena, cooling ratio records show distinct discontinuities, which, when compared against detailed station metadata records, are found to correspond to even minor changes in the station environment. Effects detected by this method are shown to escape detection by current generally accepted techniques. The existence of these microclimatic effects are a source of uncertainty in long-term temperature records, which is in addition to those presently recognized such as local and mesoscale urban development, deforestation, and irrigation.

1. Introduction

Homogeneous historical climate records are essential if we are to obtain reliable estimates of recent climate change. Numerous techniques exist to detect and correct nonclimatic inhomogeneities in historical climate records [see Peterson et al. (1998) for a review]. Such techniques can successfully identify significant inhomogeneities such as step changes due to major station relocations (e.g., large distances, elevation changes, and different settings) and network-wide changes to observational practice, or warming trends due to large-scale urbanization, but microclimatic biases that affect every climate station to some degree often escape detection. Microclimatic biases include minor instrument relocations (e.g., small distances, similar elevation, and setting), growth or removal of surrounding vegetation, encroachment by buildings, roads, and airport runways, changes in irrigation and crop practice, and even maintenance (or more often, lack thereof) of the site, instruments, and instrument shelters. While these kinds of bias are assumed to be too small to obscure true climate signals in regional- or global-scale averages, their undoubted continued presence in the global database is still cause for concern (Davey and Pielke 2005). Further, effects on individual records can be significant, potentially rendering some historical datasets effectively useless for micro- or local-scale studies. Finally, the effects of these biases (e.g., changes in shading, sky-view factor, surface radiative, thermal and moisture properties, and shelter) are likely to be pronounced in daily maximum and minimum temperatures—variables frequently used in climate change analyses.

This paper outlines the rationale and development of a new homogeneity assessment technique designed specifically to identify microclimatic inhomogeneities in historical temperature series, through examination of ratios of nocturnal cooling at neighboring climate stations. It explains why cooling ratios are an appropriate measure of microclimatic change and demonstrates that
the analysis of time series of cooling ratios, through the application of a technique called Hurst rescaling, has elucidated hitherto unrecognized inhomogeneities in several Canadian climate records.

2. Concepts and methods

a. Cooling ratios

Cooling magnitude (\(\Delta T\), °C) refers to the decrease of air temperature from the daytime maximum (\(T_{\text{max}}\)) to the minimum temperature that occurs during the subsequent night or early the next morning (\(T_{\text{min}}\)), calculated as

\[
\Delta T = T_{\text{max}(j-1)} - T_{\text{min}(j)}
\]  

where \(j\) is the day of the year. This definition is almost the same as that of diurnal temperature range (DTR), differing only in respect of the time period. The measure here extends across the night between two calendar days, hence the focus is on the cooling limb of the diurnal temperature pattern (typically between \(T_{\text{max}}\) in midafternoon and \(T_{\min}\) just after sunrise the next day).

Historically, most climate stations have used standard maximum and minimum thermometers observed once or twice per day. It is acknowledged that certain limitations of standard maximum and minimum thermometers and/or observational procedures (WMO 1996, Part I, chapters 1 and 2), may lead to occasional errors in the observed temperatures and therefore in the calculated cooling magnitude. Using these instruments, there is no way to determine the exact time the observed maximum and minimum temperatures occurred, therefore it is possible in some cases, that \([T_{\text{max}(j-1)} - T_{\text{min}(j)}]\) may not represent the actual nocturnal temperature decrease. For example, sudden changes in airmass type or cloud may disrupt the normal diurnal course of air temperature, causing the maximum and minimum temperatures to occur at other than the normally expected times of the maximum and minimum. In addition, Bootsma (1976) demonstrated that if the daily minimum temperature observation is taken, and the thermometer is reset close to the time that the minimum temperature occurs, there is an increased probability that the same temperature will be recorded on two consecutive days if the second day is warmer. It is assumed, however, that in the majority of cases \(\Delta T\) is a reasonable estimate of nocturnal cooling.

Daily cooling ratios \((R)\) of a pair of neighboring climate stations are calculated by dividing the cooling magnitude, \(\Delta T\), of a test station \((A)\) by that of a nearby reference station \((B)\):

\[
R_{A:B} = \frac{\Delta T_A}{\Delta T_B}. 
\]  

b. The weather factor, \(\Phi_w\)

Microclimatic controls on nocturnal cooling are most pronounced during calm, cloudless weather when turbulent mixing and advection are small and surface long-wave radiative losses are maximized. Under such conditions, temperature and cooling differences between two sites are greatest, so one anticipates that cooling ratios will also show significant variation with weather. However, upon examination, cooling ratios appear to be almost invariant with weather in the mean. This interesting characteristic makes them a useful indicator of microclimatic differences under all weather conditions.

Here the effects of cloud amount and type, and wind speed, on the strength of nocturnal cooling are combined into a single weather factor, \(\Phi_w\), following Oke (1998):

\[
\Phi_w = (1 - km^2)u^{-1/2},
\]  

where \(k\) is the Bolz (1949) correction factor for cloud height [i.e., \(k\) accounts for the decrease of cloud-base temperature with increasing cloud height, and ranges from 0.16 for cirrus to 1.0 for fog; for values, see Sellers...
(1965) or Oke (1987)], \(m\) is cloud amount (tenths), and \(u\) is mean wind speed (m s\(^{-1}\)). Values of \(\Phi_n\) range from zero (poor cooling conditions: overcast low cloud and/or strong winds) to 1 (excellent cooling potential: clear skies and wind speed < 1 m s\(^{-1}\)). Note a practical lower limit of 1 m s\(^{-1}\) is specified in order to avoid unreasonably large values of \(\Phi_n\) when \(u^{1/2}\) is < 1. It also ensures a convenient maximum value of 1.0.

This weather factor has been shown to be related to urban–rural temperature differences (i.e., the urban heat island effect; Oke 1998; Runnalls and Oke 2000), and potentially useful in the context of cooling and cooling ratios.

c. Hurst rescaling

In his analysis of historical hydrological records related to flooding of the Nile River, Harold Edwin Hurst developed a technique to transform time series that identifies long-term persistence in geophysical records (Hurst 1951). Subsequently, Hurst rescaling has been successfully used to identify distinct physical régimes within a variety of geophysical time series (Mandelbrot and Wallis 1969a; Outcalt et al. 1997) and fractal geometry (Mandelbrot 1982).

Hurst rescaling entails subtracting the mean of the time series from each of the original observations. The resulting deviations are then summed cumulatively to create a transformed time series, \(Q_i\). Thus for time series of monthly median cooling ratios, \(R_i\), for \(n\) months:

\[
Q_i = \sum_{i=1}^{n} (R_i - \bar{R}),
\]

then,

\[
r_n = Q_{\text{max}} - Q_{\text{min}}
\]

where \(r_n\) is the adjusted range of the series, and \(Q_{\text{max}}\) and \(Q_{\text{min}}\) are the maximum and minimum values of \(Q_i\), respectively. The rescaled range is \(r_n/s\), where \(s\) is the standard deviation of the original time series (Mandelbrot and Wallis 1969b). The rescaled range should increase asymptotically with the square root of the number of observations, \(n\),

\[
r_n/s \approx nh^n,
\]

and \(H\), the Hurst exponent, can be approximated by (Outcalt et al. 1997),

\[
H = \log(r_n/s)/\log n.
\]

Here, \(H\) is simply related to the fractal dimension \(D\) by \(D = 2 - H\) (Voss 1988). The Hurst exponent has an expected value of 0.5 for a stationary series, indicating the cumulative series is a random walk; uncorrelated white noise. However, many geophysical time series have a value of \(H\) between about 0.6 and 0.9 (Outcalt et al. 1997)—an occurrence termed the Hurst phenomenon. Such elevated values of \(H\) are an indication of persistence, nonstationarity in the record mean, or of having pooled samples with different distribution characteristics (Outcalt et al. 1997); \(H\) quantifies the nonlinear dynamics in the system.

If the accumulated deviations from the mean are normalized by the adjusted range of the series \((Q_i - Q_{\text{min}})/r_n\), the normalized series can be plotted and visually inspected to identify distinct physical régimes (Outcalt et al. 1997). Régime transitions are easily identified as inflection points in the normalized series. For example, a persistent period of above-average conditions transforms into an ascending trace, because positive differences from the mean are accumulating. A transition to below-average values is marked by an inflection point (i.e., a change from a positive to a negative slope). Steeper slopes result from larger deviations from the mean. Hence, changes in either the magnitude or sign of the slope signal régime transitions. If the value of \(H\) for a series equals 0.5, then inflection points on the rescaled trace are not significant; that is, the series are considered homogeneous. However, if \(H\) exceeds 0.5, the rescaled trace should be inspected for inflections in the trace. There is no simple indication as to what deviation from 0.5 constitutes evidence of persistence only that its strength increases as \(H\) approaches unity.

There are several advantages to this rescaling technique, the most obvious being ease of interpretation compared to the original, often highly variable time series, also it eliminates the need to arbitrarily select periods to test within records (Outcalt et al. 1997). Here we use Hurst rescaling to analyze time series of cooling ratios \(R_i\) to see if they identify microclimatic régimes. Changes in the physical controls of the microclimatic environment of a site should affect nocturnal cooling, and become evident as régime transitions when the normalized transformed series is plotted. Station history files, or metadata (Aguilar et al. 2003), can be used to see if inflection points correspond to documented physical changes to the microenvironment, of either station in the pair. If metadata records are not available to help identify the cause of the inflection point, it is still reasonable to interpret change points in \(R\) time series as inhomogeneities capable of biasing the record.

3. The nature of cooling ratios

Thermal microclimate characteristics, including \(\Delta T\), are the outcome of the unique mix of radiative, thermal, moisture, and roughness properties of the surface at a station in combination with its particular landscape
setting and external forcing by radiation, wind, and precipitation as expressed by its surface water and energy balances (Oke 1987).

a. Premises underlying the utility of cooling ratios

The usefulness of $R$ in demonstrating microclimatic effects in temperature records is based on four premises:

- Neighboring climate stations are subject to the same local-, meso-, regional-, and synoptic-scale controls, so differences in cooling between two such stations (as expressed by $R \neq 1$) should only reflect site-specific microclimate differences.
- If the microclimates of two stations do not change over time $R$ remains constant. Conversely, temporal changes of $R$ indicate changes to the microclimatic environment of at least one of the stations.
- Microclimatic influences on air temperature are most evident at night. This is so because stable conditions cause the effects of surface controls to be concentrated in the surface layer, producing well-defined spatial (both vertical and horizontal) differences of temperature.
- Microclimatic controls on nocturnal cooling are seen best during ideal (calm or near calm and cloudless) weather conditions that are conducive to cooling and strong stability.

Since cooling ratios are relatively novel measures here, we illustrate the nature of $R$ to establish that they are appropriate indicators of site microclimates, and that temporal changes in $R$ are able to identify temperature inhomogeneities due to microclimatic change. Analyses of $R$ for several Canadian climate station pairs are presented to aid understanding of their behavior and characteristics.

b. Effect of weather on the variability of $R$

The following discussion is based on cooling magnitudes calculated from temperatures observed at three climate stations near Vancouver, British Columbia: Vancouver International Airport (YVR), Steveston, and Ladner.\(^1\) All three are located on the flat lands of the Fraser River delta near the Strait of Georgia. Steveston is the reference station in both cases, because it is a long-running station of good quality that has been largely unaffected by development or urban encroachment over the period of its record, except for two minor station moves (Schaefer 1974). YVR is a principal synoptic station, with both hourly and daily records. It is located approximately 5 km north of Steveston, and both are at almost equal distances from the deltaic coastline and at similar elevations (2 and 0 m ASL, respectively). The big difference between them is the fact that YVR has experienced considerable change because of the growth of the international airport and the associated construction of runways, taxiways, hangars, terminals, and parking lots. The YVR and Steveston temperature records overlap in the period from 1937 to 1970. The third station, Ladner, is approximately 12 km east-southeast of Steveston, at an elevation of 1 m ASL, and its temperature observation period overlaps with Steveston between 1953 and 1970. The Ladner station was located at a Canadian Forces Base, and its surroundings remained essentially unchanged over the period of the record. Both Steveston and Ladner are ordinary climate stations, having daily maximum and minimum temperatures and precipitation records (for photographs of these two stations see Fig. 8).

Figure 1 shows an approximately linear relation between $\Phi_w$ and the magnitude of nocturnal cooling at Steveston for January and September 1955. The weather factor $\Phi_w$ ranges from 0 (strong winds, overcast low cloud) to 1.0 (calm, clear), see Eq. (3). Linear regression $r^2$ statistics: 0.71 and 0.70 for September and January, respectively.

1 A full set of station history information for sites mentioned in this paper are available as Tables A1–A4, at http://dx.doi.org/10.1175/JCLI3663.S1.
of cooling magnitude. In general, because of the difference in soil thermal properties and plant cover, a given value of $\Phi_w$ leads to greater cooling in September than January, but in both seasons relatively greater cooling occurs in less cloudy and windy weather conditions (i.e., as $\Phi_w$ increases). The relatively good correlation ($r^2 \sim 0.7$) confirms $\Phi_w$ is an appropriate descriptor of weather conditions relevant to nocturnal cooling.

The cooling ratio can also illustrate the seasonality of cooling. Daily $R$ values for an entire year (1955) are plotted as a function of $\Phi_w$ for YVR: Steveston; and Ladner: Steveston (Fig. 2). This year was selected because August 1955 was one of the driest on record (2.8 mm precipitation). Hence, this year might be expected to represent a wide range of seasonal differences, if they are significant. The most prominent feature of Fig. 2 is the decreased scatter with increasing $\Phi_w$, for both station pairs. Stable values of $R$ emerge as $\Phi_w$ approaches unity. This value is approximately 0.75 for YVR: Steveston, and 1 for the Ladner: Steveston pair. To a first approximation, these values appear to be approximately equal to the mean for the entire series, indicating that weather variations affect the variability, but not the central tendency, of $R$. This conservative behavior is worth exploiting in climatological analysis.

Some of the variability of $R$ seen in Fig. 2 may be due to differences in the seasonal variation of microclimatic controls on cooling at the individual sites (e.g., soil moisture, vegetation growth, snow cover, etc.). Closer examination of occasions with low $\Phi_w$ shows that many $R$ values approach unity when mean wind speeds exceed 8 m s$^{-1}$. This is not surprising since strong winds eliminate spatial variations in temperature due to strong mixing and advection. However, low values of $\Phi_w$ that arise from a combination of light winds and overcast, low clouds still permit some expression of site control, because $R/\Phi_w$ are observed under these conditions.

A plot of $R_{YVR\text{-}Steveston}$ against the absolute cooling magnitude ($\Delta T$) at Steveston (Fig. 3), reveals that most of the large scatter occurs when the absolute cooling magnitude itself is small (6°C or less). For example, a ratio of $1°C/0.5°C = 2$, but is essentially meaningless because actual cooling was insignificant. Extreme outliers can also arise as a result of data errors. Clearly, some form of data screening is warranted to eliminate meaningless outliers and occasions when significant cooling does not occur. An optimal method to accomplish this is discussed in section 3f.

We conclude that relative nocturnal cooling differences between sites are preserved under less than ideal weather conditions, even when temperature differences themselves may be quite small and hence $R$ for a station pair is a characteristic measure of microclimate difference. This difference is more easily identified under ideal weather conditions (when the scatter is reduced), but is also evident in the mean when conditions are less than ideal.

c. Seasonal variations of cooling ratios

The possibility that some of the scatter apparent in Fig. 2 is due to seasonal influences is confirmed by Fig. 4. The 1955 monthly median $R$ (i.e., median value of daily $R$ for each month) for nights with $\Phi_w > 0.5$ at the same two Vancouver area station pairs indicate that, while both series contain seasonal variations, the relation between the two is antiphase (i.e., maximum
 occurs with minimum \( R_{YVR:Steveston} \) and vice versa). This pattern is explained by the seasonality of soil moisture differences between the sites. The seasonality of \( R_{YVR:Steveston} \) is consistent with typical variations in soil moisture in the area, and the associated variation in thermal properties (Runnalls and Oke 2000). The pattern is one of wet, thermally damped soils in winter and early spring and dry, thermally "flashy" soils in summer and autumn. One of the main differences between YVR and Steveston is the presence of extensive tarmac surfaces at YVR (runways, taxiways, parking lots). Greater differentiation between the two microclimatic environments is expected in summer when soils are dry, because the moist and wet soils found in winter are thermally similar to tarmac. August 1955 was one of the driest on record, hence it is consistent that cooling at YVR was only 73% of that at Steveston. Values of \( R_{Ladner:Steveston} \) close to unity imply similar cooling regimes at both sites, which is consistent with their similar open, rural exposure, soil moisture, and soil thermal properties.

The seasonality of \( R \) is likely to vary with the specific climatic régime of an area. The seasonal variations near Vancouver are mainly related to soil moisture; elsewhere they may be controlled by changes in snow cover, vegetation phenology, agricultural practices, or artificial heating.

It is possible that the seasonal pattern of \( R \) could change over time, indicating microclimatic change, or other inhomogeneities at one of the stations. Figure 5 shows a clear example of this possibility. The monthly median \( R \) for the station pair of Regina International Airport: Regina Canada Department of Agriculture (CDA) is plotted for different time periods. Both sites are on flat land: CDA is an agricultural research farm of long standing at a truly open rural site; Regina Airport has seen significant expansion of airport development (Fig. 11). The seasonal pattern for the later period (1984–93) is the mirror image of the earlier period (1932–41), while the average calculated for the entire series (1932–93) displays no seasonal variation, that is, the two patterns cancel each other. In this particular case, the change in the seasonal pattern is the result of a discontinuity in winter minimum temperatures at Regina CDA, which is apparent when examining the temperatures alone (Gullett et al. 1990). The station history file does not provide evidence for the cause of this step change. No conclusions can be drawn as to whether it is indeed caused by a feature of the microclimatic environment, or perhaps as a result of a change in instrumentation or observation procedure. However, because the absolute value of the summer \( R_{Regina Airport:Regina CDA} \) value also decreased, without a corresponding discontinuity in either the maximum or minimum temperature series, there may be more than one cause of the changes. This example is included to demonstrate the importance of the time period selected for analysis. In this case, if the entire length of record had been used to examine seasonality, one might erroneously conclude that \( R \) for this station pair has no seasonality.

**d. Soil moisture changes and variability of \( R \)**

As noted in the preceding section, seasonal variations of the \( R_{YVR:Steveston} \) are probably linked to variations of soil moisture (thermal properties). Therefore, it must be considered whether soil moisture variations from year to year in a given season could also induce changes to \( R \); that is, do exceptionally wet summers lead to more winter like \( R \) values? This question was examined by comparing \( R_{YVR:Steveston} \) for the three driest Augs (average monthly precipitation 7.8 mm) and three wettest Augs (average monthly precipitation 77 mm). Median \( R \) for both groups was 0.81. A plot of median monthly \( R \) for both station pairs against total
monthly precipitation for the period 1953–60 shows no strong relation. Analysis of daily $R$ values as a function of antecedent precipitation similarly shows no relation ($R$ is the same at 0 and 80 mm precipitation; Fig. 6a). Therefore, it appears that the range of summertime soil moisture in the Vancouver region does not significantly affect the estimation of $R$. This independence may not hold in other regions, but because most climate stations also record precipitation, the possibility of soil moisture variations causing temporal variations of $R$ can be investigated on a site-specific basis.

e. Relation between absolute temperature and $R$

A difficulty in the detection of inhomogeneities in temperature series arises when site-specific trends are superimposed on regional temperature trends. It becomes difficult to determine how much of the observed temperature trend is due to regional warming or cooling, and how much is a site-specific effect.

Given the relative stability of $R$ with variations in weather, it seems likely that $R$ would also remain constant even if regional trends of temperature exist, because effects at the regional scale are common to both stations. On the other hand, it also seems possible that microclimatic differences are accentuated in warmer periods if differences of surface thermal properties are important to the cooling ratio. That is, during warmer (probably summer) periods, daytime heat storage may be enhanced and site-specific storage differences may lead to greater nocturnal cooling differences. Therefore, we need to consider whether temporal trends in $R$ occur as a result of regional warming or cooling.

Zhang et al. (2000) report that southwestern British Columbia experienced a statistically significant increase in minimum daily temperatures of approximately $1^\circ$ to $1.5^\circ$C in all seasons between 1900 and 1998. Over the same period, the region also experienced a statistically significant decrease of approximately $0.5^\circ$ to $1.0^\circ$C in diurnal temperature range. Zhang et al. (2000) also show that similar, but slightly smaller, trends are evident in the period from 1950 to 1998 as well, although the trends in diurnal temperature range were not statistically significant during the period. The time series of monthly $R_{Ladner:Steveston}$ (1953–70) does not exhibit trends (Fig. 7b). The fact that the $R$ remains remarkably constant during this period of regional warming shows that the use of a ratio eliminates the regional influence common to both series. Again, this argues that $R$ is a useful indicator of microclimatic differences between sites.
The possibility that regional warming trends might affect time series of \( R \) is considered by examining the relation between median \( R \) in August and monthly mean maximum temperatures observed at Steveston to determine if \( R \) values were significantly different in warmer than colder years. As Fig. 8 shows, there is no trend in \( R \) even though mean maximum temperatures range from 20° to 24°C. Again, this suggests that regional warming or cooling trends are unlikely to introduce temporal trends in time series of \( R \), rendering it a useful indicator of microclimatic differences.

f. Optimal estimation of \( R \)

The foregoing analyses indicate that daily \( R \) for a climate station pair exhibits apparently random scatter not explained by weather, temperature, or soil moisture variations, however, the central tendency of \( R \) remains constant over time, except for seasonal cycles. It is reasonable to conclude that a measure of the central tendency of \( R \) for a given pair of climate stations is a characteristic feature of the station pair at a given time that represents microclimatically controlled cooling differences between the sites. This characteristic remains constant, even with regional climate changes (i.e., trends in weather patterns, temperatures, or precipitation). Given the relative stability of \( R \) for a station pair, this measure seems potentially useful as a way to identify site-specific microclimatic biases in temperature records. However, given the day-to-day variability, it is necessary to first identify an optimal estimate of \( R \).

Monthly or seasonal averages can obscure microclimatic biases that are more pronounced on certain days, because of favorable weather and surface conditions, than on others. Therefore, while it is desirable to retain the information content present in daily data, some measure of central tendency is needed because of the scatter present in such data. Figure 2 shows that the scatter decreases as cloud and wind decrease, and that the characteristic \( R \) becomes apparent at higher values of \( \Phi_w \). In keeping with the initial premise that microclimatic effects are greatest under ideal weather conditions, it is desirable to calculate \( R \) for only these conditions. On the other hand, Fig. 2 also shows that the central tendency of \( R \) does not appear to vary significantly with weather, only that scatter is greater and outliers are more common at low \( \Phi_w \). Therefore we need to know if including the scatter biases the estimate of central tendency compared to using only \( R \) for ideal weather situations, and whether the extra input data requirements and computations needed to calculate daily \( \Phi_w \) are warranted.

Table 1 summarizes statistics estimated from several subsamples of August \( R_{YVR:Steveston} \) values for 1953–60. This period is chosen to maximize the sample size, while at the same time ensuring that temporal trends are not present in the \( R \) series (a shift upward in \( R \) occurs around 1962). In addition to statistics for the entire dataset (1953–60), Table 1 also includes three subsets of weather conditions for cooling: poor weather only (\( \Phi_w \leq 0.5 \)); favorable weather only (\( \Phi_w > 0.5 \)); and the 50th percentile of cooling events (i.e., occasions when cooling at Steveston was relatively large). The
last subset is included to determine if it is possible to select good cooling events without the need to calculate \( \Phi_w \) for every day. Support for this alternative would make the technique more widely applicable, because many climate stations do not have the hourly synoptic data from a station nearby needed to calculate \( \Phi_w \).

Results indicate that the scatter at low \( \Phi_w \) does bias the estimate of the mean and the median. The \( R \) values for poor weather conditions are significantly higher (95% confidence limit) than those with favorable weather. Inclusion of all weather conditions slightly overestimates the mean and median, relative to the \( \Phi_w > 0.5 \) set. There is good agreement of the means and medians estimated from the \( \Phi_w > 0.5 \) set, and the 50th percentile cooling data. Using the latter population increases the sample size by 75%, but ensures nights that did not experience significant cooling are excluded. These results suggest that estimating statistical properties from best cooling events in a month (the 50th percentile) appears to be the optimal approach. This method has the considerable added benefit that reliable determination of \( R \) can be obtained without the need for detailed ancillary weather data.

Figure 7 shows time series of daily August \( R_{YVR:Steveston} \) and \( R_{Ladner:Steveston} \) values for 1953 to 1970. Day-to-day variability is apparent, but the mean value is constant throughout the whole period in the Ladner: Steveston case, but there is a slight shift upward in the mean signal for YVR: Steveston around 1962. The timing of this shift corresponds to station moves at both Steveston and the airport.

g. Potential advantages of cooling ratios for homogeneity analysis

We conclude that the cooling ratios defined here are potentially useful to identify microclimatic biases in temperature records. They offer several advantages over existing methods of homogeneity analysis, namely:

- cooling ratios derived from occasions when strong cooling occurs, and therefore, when microclimatic bias should be most pronounced, reduce the possibility that subtle biases are obscured, as they are when incorporated in monthly, seasonal, or annual averages.
- While temperatures themselves are inherently variable because they respond to regional, synoptic, or global climate trends, cooling ratios appear to be remarkably stable (i.e., invariant with absolute temperature or soil moisture), making them better suited to the detection of small, site-specific effects and trends.
- Here \( R \) is dimensionless, which simplifies interpretation and permits comparison between sites.
- Cooling ratios are calculated from daily maximum and minimum temperatures, which are routinely observed at almost every climate station, so the technique is widely applicable as long as a regional reference station is available.

Nocturnal cooling is a physically relevant process-driven phenomenon related to the microclimatic environment of climate stations. Cooling ratios are therefore likely to be a physically meaningful measure of the degree of thermal modification of the surroundings of a climate station. This is likely to be distinctly superior to surrogate measures of the degree of anthropogenic site disturbance such as population in the surrounding census tract (Karl et al. 1988; Jones et al. 1990), or satellite-derived night-light intensities (Hansen et al. 2001) or normalized difference vegetation index (NDVI; Gallo and Owen 1999), which are commonly used to designate stations as either rural or urban in attempts to study long-term temperature trends free of the effects of urban development. Analysis of cooling ratio time series is potentially useful as a tool to identify site-specific microclimatic biases in temperature series.

4. Application of the Hurst rescaling technique

Here Hurst rescaling is applied to time series of \( R \) from several climate station pairs. In the following examples, the original time series are monthly median \( R \) values for the 50th percentile of cooling events (i.e.,

---

**Table 1. Comparison of statistics for subsets of August cooling ratios, \( R_{YVR:Steveston} \) for the period 1953–60.**

<table>
<thead>
<tr>
<th></th>
<th>All data</th>
<th>( \Phi_w \leq 0.5 )</th>
<th>( \Phi_w &gt; 0.5 )</th>
<th>50th percentile*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum ( R )</td>
<td>2</td>
<td>2</td>
<td>1.08</td>
<td>1.04</td>
</tr>
<tr>
<td>Minimum ( R )</td>
<td>0.45</td>
<td>0.45</td>
<td>0.59</td>
<td>0.45</td>
</tr>
<tr>
<td>Mean ( R )</td>
<td>0.79</td>
<td>0.82</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>Median ( R )</td>
<td>0.77</td>
<td>0.81</td>
<td>0.73</td>
<td>0.72</td>
</tr>
<tr>
<td>Standard deviation of ( R )</td>
<td>0.16</td>
<td>0.18</td>
<td>0.10</td>
<td>0.09</td>
</tr>
<tr>
<td>N</td>
<td>248</td>
<td>179</td>
<td>69</td>
<td>122</td>
</tr>
<tr>
<td>95% confidence limits for mean, median</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

* Above-average cooling at Steveston.
only nights with above-average cooling are used). Series are rescaled and the Hurst exponent is estimated. The station history files used in the analyses are periodic reports by inspectors including photographs, site plans, maps and miscellaneous metadata held in the archives of the Meteorological Service of Canada. Station records were selected based on several criteria: the detail and completeness of the station history file; length and completeness of the temperature records; availability of and proximity to other stations in order to calculate meaningful cooling ratios; uniformity of regional characteristics to minimize other sources of between-station differences; and the presence of site-specific features or changes that might illustrate the technique.

**Example 1: Homogeneous rural series**

The first example of Hurst rescaled $R$ is the simple rural station pair of Ladner and Steveston, near Vancouver, British Columbia. Statistics for the original and the rescaled series are given in Table 2. Temperature data are available for the 1953–70 period. This example illustrates the importance of calculating the Hurst exponent prior to interpreting the rescaled trace. In this case, the estimated value of $H$ is 0.5, which indicates the absence of mixed régimes; essentially it is a random series. The apparent stationarity of the mean $R$ supports this conclusion, however the Hurst rescaled trace shows a number of inflection points, which could be misinterpreted as significant régime transitions, if the Hurst exponent had not been estimated. This highlights the importance of estimating $H$ and identifying inflections in the rescaled time series—both are required.

According to the available metadata (Table A1 in electronic supplement to this article, available at http://dx.doi.org/10.1175/JCLI3663.S1), there is reason to expect the individual temperature series to be homogeneous, and thus $R$ to remain constant over time. Both

**Example 2: Multiple inhomogeneities: Airports**

Edmonton Municipal Airport (EMA) and Edmonton International Airport (EIA) are both first-order synoptic stations near Edmonton, Alberta (Figs. 9a,b). EMA lies within the city of Edmonton, (3 km northwest of the city center), and is known to be influenced by the urban heat island effect of the city (Hage 1972). In addition, Gullett et al. (1991) found the daily minimum temperature exhibits an inhomogeneity, a warming trend from 1968 to 1983, which is consistent with the expected influence of a growing urban heat island. EIA, on the other hand, is 26 km south-southwest of the city center, and is considered outside of the city’s thermal influence (Hage 1972). The temperature records for the two stations overlap from 1961 to 1998. The stations are attractive here because of the availability of extensive station history for both stations, which should aid interpretation of the significance of inflection points in the rescaled time series. The Hurst exponent for the $R$ series was 0.76 (Table 2) indicating the presence of a mixture of cooling régimes, which should be apparent in the rescaled trace.

### Table 2. Summary of statistics for time series of $R$, and Hurst rescaled values. For definition of symbols see text following Eqs. (4) and (5).

<table>
<thead>
<tr>
<th>Time series</th>
<th>$n$</th>
<th>Mean</th>
<th>$s$</th>
<th>$Q_{\text{max}}$</th>
<th>$Q_{\text{min}}$</th>
<th>$r_n$</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ladner: Steveston</td>
<td>193</td>
<td>1.06</td>
<td>0.07</td>
<td>0.40</td>
<td>−0.57</td>
<td>0.97</td>
<td>0.50</td>
</tr>
<tr>
<td>Edmonton Municipal Airport:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edmonton International Airport</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly</td>
<td>456</td>
<td>0.83</td>
<td>0.07</td>
<td>3.38</td>
<td>−4.13</td>
<td>7.51</td>
<td>0.76</td>
</tr>
<tr>
<td>Annual</td>
<td>38</td>
<td>0.83</td>
<td>0.04</td>
<td>0.27</td>
<td>−0.34</td>
<td>0.61</td>
<td>0.73</td>
</tr>
<tr>
<td>Woodbend: Ellerslie</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1973–86</td>
<td>153</td>
<td>1.15</td>
<td>0.08</td>
<td>0.08</td>
<td>−2.48</td>
<td>2.55</td>
<td>0.69</td>
</tr>
<tr>
<td>1979–86</td>
<td>92</td>
<td>1.17</td>
<td>0.26</td>
<td>1.14</td>
<td>−0.04</td>
<td>1.19</td>
<td>0.34</td>
</tr>
<tr>
<td>Regina Airport: Midale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1923–91</td>
<td>818</td>
<td>0.99</td>
<td>0.10</td>
<td>0.23</td>
<td>−13.29</td>
<td>13.51</td>
<td>0.73</td>
</tr>
<tr>
<td>1963–91</td>
<td>346</td>
<td>1.03</td>
<td>0.11</td>
<td>3.94</td>
<td>−0.64</td>
<td>4.57</td>
<td>0.64</td>
</tr>
<tr>
<td>Regina Airport: Area composite</td>
<td>714</td>
<td>1.06</td>
<td>0.09</td>
<td>1.83</td>
<td>−12.02</td>
<td>13.80</td>
<td>0.77</td>
</tr>
</tbody>
</table>

### Effect of Seasonal Variations of $R$

Outcalt et al. (1997) showed that seasonal cycles do not need to be removed before applying the technique, because they do not change the overall pattern of the
rescaled trace. This appears to be the case for rescaled $R$ time series. For example, Fig. 9c shows the rescaled $R$ for the annual means at EMA and EIA. If the equivalent trace for the monthly means is added (not shown) the two curves are almost identical with only small differences and no changes in the location of major inflection points. For the annual series, $H = 0.73$, compared to 0.76 for the monthly series (Table 2). In general the use of monthly $R$ values is preferred. While annual results may be slightly easier to interpret, monthly data allow identification of a specific month in which a régime transition occurred and for the annual average, the sample size is reduced by a factor of 12, which may be problematic if the original time series is short. This difficulty becomes apparent when calculating the Hurst exponent because values of $H < 0.5$ result from an insufficient sample size.

2) IDENTIFICATION OF COOLING RÉGIMES IN THE RESCALED TRACE

Figure 9c shows the $R$ time series and the rescaled trace for the two Edmonton Airports for the period 1961–91. Events noted in the station history files that could affect temperature observations are given symbols on the trace. Changes to the immediate environment of both sites are evident in the series of site photographs included in the station history files.

As an aside, we note some common difficulties encountered, even when dealing with some of the better station history files. First, it is not always possible to determine the exact date of documented events or changes. For example, inspections frequently occur a few months after a station move, so the date of the move and the inspection are not the same. At the time of an inspection, recommendations are often made, such as to paint the screen, or change the observation schedule, but that does not necessarily mean the changes occur soon, or even at all. Second, the inspectors’ reports often refer to declining exposure over a period of years, so the possible microclimatic biases tend to occur gradually. Third, the biggest problem, by far, is the lack of detail in the written reports. For example, several consecutive reports may indicate no changes in site exposure, yet photographs clearly show vegetation growth, parking lot expansion, or the construction or removal of buildings. Photographs are probably the most valuable historical evidence to document changes to the immediate environment, but often
they are taken only once or twice in the life of the station. In summary, it is unrealistic to expect exact correspondence between inflection points in rescaled series and specific dates noted in the history file, or dates inferred from the available evidence. Nevertheless, station history files remain the only source of information to explain temporal changes in the $R$ series.

Despite these limitations, there is excellent correspondence between major inflection points and events from the station histories in Fig. 9c. The rescaled trace exhibits significant régime transitions around 1965 (EIA moved), 1975 (EMA moved), 1978 (EIA moved November 1977), 1980 (parking around instruments at EIA ends), 1984 (EMA moved to roof), 1992 (no evidence available), and 1994 (EMA automated). Smaller inflection points may be due to the seasonality of $R$ values, and some appear to be associated with minor events such as painting the screen (1988).

Identification of distinct change points in the $R$ time series is complicated by the fact that both stations were affected by significant development and encroachment of the instrument sites at several times, both were particularly affected during the 1970s and early 1980s. Although the rescaled trace declines over the same period, the cooling ratios themselves show a definite upward trend (i.e., approaching a value of unity), implying that cooling régimes at the two airports were becoming more similar. This transition could be the result of changes at either station, but, given the nature of the development, a reasonable interpretation is that Edmonton International was becoming more like the urban municipal airport. Despite the complex influences on this particular time series, through the use of Hurst rescaling it is possible to identify distinct cooling régime transitions that correspond to documented changes to the microclimate environment of the individual stations.

c. Example 3: Vegetation growth

The station pair of Ellerslie and Woodbend, Alberta, illustrates the effect of vegetation growth on the homogeneity of temperature records. Ellerslie is an agricultural research center located within the city limits of Edmonton, but south of the main urban area (Fig. 10a). Although Ellerslie’s temperature record unfortunately ended in 1986, it is a good rural reference station that remained in the same location throughout its history.

Woodbend is also outside of the urban center, approximately 22 km southwest of the city (west of Ellerslie and north of Edmonton International Airport). Photographs show that when Woodbend was first established, the surrounding area had been cleared of trees (Fig. 10b). Subsequent photographs show the trees grew back over time providing significant shelter to the instrument area (Fig. 10c). Woodbend is an excellent example of vegetation growth around a climate station without other confounding changes over time (i.e., construction, station moves, observer changes, etc.). Whereas Woodbend’s temperature record extends from 1973 to present, Ellerslie’s extends from 1963 to 1986, providing a 13-yr record overlap between the two.

Figure 10d shows the original $R$ series for Woodbend: Ellerslie, as well as the rescaled trace, and the rescaled trace determined for the period 1979–86. A step change in the original $R$ trace is apparent around 1979, and this inflection is clearly supported by the change in slope of the rescaled trace. For this reason, a second Hurst rescaling analysis was conducted for the post-1979 period to determine if the large step change might obscure smaller inflection points in the trace. Comparison of the two rescaled traces indicates that while quite different in form, trace inflections occur at the same time in both series, implying that the 1979 step change does not prevent the identification of smaller inflection points. Furthermore, the Hurst exponent was estimated for both series, with $H = 0.69$ for the longer record, and $H = 0.34$ for the shorter record (Table 2), implying that the shorter record is less reliable because of its smaller sample size.

The step change in 1979 is attributed to the Woodbend rather than the Ellerslie series, because a similar step change occurs in the $R$ series for Woodbend: Edmonton International (not shown). The station file indicates that late in 1978 the observation schedule at Woodbend changed from twice daily to only once per day. Such changes in observation schedules have detectable effects on series of minimum temperatures (Bootsma 1976). Interestingly, the 1978 inspector’s report for Woodbend notes “no changes” to the surrounding environment, whereas every report after that (starting in 1981) notes increased tree height. Therefore, perhaps at least part of the 1979 change in the $R$ series is attributable to the noticeable growth of vegetation.

Upon examining the period following the step change in 1979, a small negative trend in the original $R$ series is identifiable; mean cooling at Woodbend decreases 7% relative to the more open Ellerslie site over the 8-yr period. This is consistent with the reduced radiative cooling to be expected because the sky-view factor decreases as the horizon starts to become screened by the growth of the trees. However, note that vegetation growth can have other effects on the cooling régime as well, such as preventing cold air from draining away from the site and moving the active surface higher as the canopy layer deepens. Vegetation growth
could also influence the daytime maximum temperature in several ways: cooling the ground by increased shade, increased evapotranspiration, and reduced natural ventilation of the Stevenson screen. Thus, while Hurst rescaling appears useful to identify physical cooling régime changes, the task of determining the exact physical mechanism responsible for the change is not straightforward.

Despite the fact the Hurst exponent for the 1979–86 rescaled \( R \) series is small, which implies the sample is too small to allow reliable interpretation of inflection points, most of the series breaks do correspond with changes documented in the history file, such as the replacement of the screen in 1981, and the planting of additional trees near the instruments in 1983.

d. Example 4: Rural airport

1) Using a single reference series

Cooling ratios for the station pair Regina Airport and Midale, Saskatchewan, were of interest for two reasons: first, the flat, uniform regional topography minimizes potential sources of between-station differences, ensuring that any differences are likely due to site-specific influences; and second, maximum and minimum temperature series from both stations have been analyzed previously by Gullett et al. (1990, 1991) who found both records to be homogeneous despite station history information from Regina Airport indicating significant encroachment by buildings and runways. Thus, comparison between the Gullett et al. (1990) method and the Hurst rescaling technique might help to establish whether the latter offers any advantages in the detection of site-specific inhomogeneities.

Midale is a climate station with a long time series located approximately 120 km southeast of Regina. Despite a few small moves, its surroundings have remained distinctly rural throughout its history (Figs. 11a,b), although for a brief period it was located in a sheltered garden. Regina Airport, on the other hand, has seen considerable development including a
Fig. 11. The Midale site in (a) 1934 and (b) 1991. The Regina Airport site over a 60 year period: (c) in 1934 viewed from the south (left) and northeast (right), (d) in 1964 viewed from the north (left) and east (right), and (e) a full panorama in 1993. Note the station site moved around in the same general area. A new terminal was built in 1960 and the station moved; it was moved again, further south, in 1983; a large new building was constructed to the north in 1989, and new buildings to the east and in the airport environs were built by 1993.
new terminal building, aircraft taxiways, and car parking in 1960 (Fig. 11d), later extensions to the terminal, and the addition of new parking lots, taxiways, and buildings in and around the airport (Fig. 11e). Further, the station has been moved around the area especially in 1983. Finally, inspector reports throughout the 1980s repeatedly refer to the heat island problem at the airport site.

It is helpful to examine the $R$ series and the rescaled trace together (Fig. 12a) in order to interpret the changes identified. The fact that $H = 0.73$ signals the existence of nonlinear effects. Because of a shift in the $R$ series in about 1961 there is a major inflection point in the rescaled series (Fig. 12a), so rescaling analysis was also performed on the period 1963–91 (Fig. 12b) to make smaller régime transitions easier to identify. Statistics for both time series are given in Table 2. The probable cause of the change in 1961 is the construction of a large parking lot and new terminal building in 1960 and relocation of the instruments at around the same time.

The $R$ series (Fig. 12a, top) exhibit an increasing trend between approximately 1950 and the mid 1960s, followed by a decreasing trend in the annual mean of approximately 8% to 1984. The decrease is consistent with increased development occurring at Regina Airport over the same period. Instruments were moved to a more open site in the fall of 1983, and a minor inflection point in the rescaled trace is evident at about the same time. Another significant change in the rescaled trace occurs in 1988 (Fig. 12a), but the cause of this change is difficult to identify from information in the station file. Significant development occurred at Regina Airport around this time, so perhaps not all changes were documented.

Other inflection points in the rescaled trace (Fig. 12a, lower) correspond to the station move from the City of Regina to the airport in 1932, and substantial construction (runways and buildings) in the late 1930s. A slight change in slope of the rescaled trace occurs after the relocation of the Midale station in 1964. Two other relocations of the Midale station in the 1950s do not appear to affect the rescaled trace, nor do other noted events such as painting the screens. There is no evidence in the station history files to explain the inflection points around 1927, 1931, and 1980.

2) Using a Composite Reference Series

As the preceding analyses show, temperature records from stations that appear homogeneous may still exhibit discontinuities due to subtle site changes or instrument changes, etc. Thus, when time series of $R$ using a single reference station are constructed, it is difficult to know whether discontinuities in the series are due to the test, or the reference station. If sufficient station history information is not available, as is often the case, it may be impossible to determine the source of the discontinuity. This problem is addressed in other types of homogeneity analysis by using composite reference series constructed from the records of a number of surrounding stations (e.g., Gullett et al. 1991; Peterson and Easterling 1994). This approach assumes that inhomogeneities in the individual stations are smoothed out, thereby reducing the reliance on any individual station and its station history information. Figure 13 shows the

![Fig. 12. (a) The $R$ and Hurst rescaled $R$ (more regular line with symbols) for Regina Airport: Midale ($H = 0.73$). Symbols denote events documented in station history files: (+ Regina Airport; O Midale). (b) Rescaled $R$ for the same station pair but just for the 1963–91 period ($H = 0.64$). Symbols same as (a).](image1)

![Fig. 13. The $R$ and rescaled $R$ (more regular line with symbols) for Regina Airport: Area composite ($H = 0.77$). Symbols (+) denote events documented in the Regina Airport station history file.](image2)
and rescaled $R$ time series for Regina Airport and an area composite reference series constructed by averaging the cooling magnitudes at four nearby stations (Midale, Lumsden, Muenster, and Regina CDA) selected to obtain the longest possible record (1932–91). The form of the resulting rescaled series is very similar to that for Regina Airport: Midale (Fig. 12a). There are major inflection points in 1939 and 1961, and some smaller slope changes in 1948 and 1989 that correspond to documented station history changes. The 1961 discontinuity in the cooling ratio series (Fig. 13, upper) is more obvious than in the top of Fig. 12a, with an apparent step change in the mean, as well as increased variability occurring around that time.

The similarity between the results using single and area composite reference series suggests the latter is a useful approach when a single suitable homogenous reference station is not available. However, particularly when subtle siting biases are the focus of investigation, the use of a single reference station with good metadata is preferable. The availability of data, metadata, and suitable reference stations therefore determine which approach to use in any particular analysis.

5. Comparison of Hurst rescaling with other techniques of homogeneity analysis

a. Double mass analysis and parallel cumulative sums

There are several techniques to detect inhomogeneities in climate records (Peterson et al. 1998). Theoretically, several of these techniques are applicable to time series of cooling magnitudes, or cooling ratios, as an alternative to Hurst rescaling. Two that potentially might yield similar results to Hurst rescaling are tested here: double mass analysis (e.g., Kohler 1949) and parallel cumulative sums (Rhoades and Salinger 1993).

Double mass analysis involves plotting the cumulative sum of a given variable (e.g., minimum temperature, cooling magnitude, or their respective deviations from the mean of the series) at a candidate station against that of a neighboring station. Changes of slope of the line indicate potential inhomogeneity in one of the records. In the present context, if cooling magnitudes are under study, then the slope of the double mass plot represents a cumulative cooling ratio for the station pair, and a change of slope should indicate a change in cooling, or a potential discontinuity, in one of the records. This approach was applied to the monthly cooling magnitudes at Regina Airport and Midale. Only very minor changes in slope of the mass curve are evident (Fig. 14a). Comparison of the double mass plot with the Hurst rescaling of the same records (Fig. 12a, lower) shows the latter better detects the subtle discontinuities in the cooling record. The reason for the difference probably lies in the magnitudes of the discontinuities relative to the original values. That is, when a time series of cooling ratios is constructed, the values are close to unity, whereas with double mass analysis in this case, the cumulative sum of monthly cooling magnitudes over a seventy-year period exceeded 13 000. Thus, small changes to the cooling ratio are more obvious than are changes to the slope of the cumulative cooling magnitudes mass curves. This problem can be avoided if shorter time periods are chosen for analysis; however, this requires first choosing the time periods to be analyzed subjectively and repeating the computations several times.

Double mass analysis of the cumulative deviations of cooling magnitude from the mean of the series might also make the discontinuities slightly easier to identify. But it appears that the rescaling of these deviations to values between 0 and 1 is the key to making the discontinuities stand out. Another advantage of Hurst rescaling is the fact that the time scale is retained in the plot; this makes it easier to identify the date of a discontinuity. Further, the original time series of $R$ and the Hurst rescaled trace can be plotted on the same scale, which helps to identify discontinuities.
Another limitation of double mass analysis is that if a discontinuity is indicated by a change of slope it is not possible to determine which of the two records is responsible for the discontinuity (Peterson et al. 1998). To overcome this Rhoades and Salinger (1993) developed a technique called parallel cumulative sums (CUSUMS), in which cumulative deviations from the long-term mean values from several neighboring stations are plotted. It is assumed that change points common to all of the series probably reflect regional climate changes, whereas inflection points that occur in only one of the series is probably a discontinuity unique to that particular record.

Figure 14b shows parallel CUSUMS for Regina Airport and Midale. Ideally, several neighboring series are plotted together so the discontinuity can be attributed to the correct record. Only two records are shown here for illustrative purposes. Although several change points are evident in the series, deciding which ones are common to both series and which are unique is not straightforward. The task is even more difficult when multiple series are examined together. Hurst rescaling has the advantage that only one trace needs to be examined, (although as indicated earlier, plotting the original cooling ratio series together with the rescaled trace is optimal), so that inflection points are easily identified.

b. Multiphase linear regression

Gullett et al. (1991, hereafter G91) tested individual Canadian climate stations for homogeneity using a multiphase regression technique. The method was extended by Vincent (1998) and used to select stations for inclusion in the Canadian Historical Temperature Database (CHTD). The CHTD contains the most reliable monthly temperature data for Canada and is intended for climate change analyses (Vincent and Gullett 1999). Here we compare the results of the rescaled cooling ratio analyses to those of G91 for some of the stations used as examples here, to see if the new technique provides a means to improve the homogeneity of the CHTD records.

1) Regina Airport and Midale

G91 report that both the maximum and minimum monthly temperature series for Midale, Saskatchewan, are homogeneous for the period 1923 to 1988 (1988 was the latest year tested for all records). They also find the minimum temperature series for Regina Airport to be homogeneous from 1904 to 1988, while the maximum temperature series contains step changes in 1913 and 1932, but is otherwise homogeneous. Gullett et al. (1990) report a number of statistically significant inhomogeneities (95% significance level) in the Regina Airport time series, especially one in the minimum temperature series between 1961 and 1963. Nevertheless, they deem the step changes to be climatologically insignificant, and the temperature record is considered homogeneous based at least partly on the lack of evidence in the station history file to indicate a cause for the inhomogeneities. Regina Airport was consequently selected to be one of 210 stations included in the CHTD (Vincent and Gullett 1999). Midale was not included in the CHTD, because it closed in 1991.

The Hurst rescaled $R$ results for Regina Airport: Midale cast doubt on the conclusion of climatological insignificance; régime transitions are apparent, which suggest that the temperature series is not homogeneous. Examination of the station history files at least implies potential reasons for some of the inhomogeneities (or cooling régime transitions). For example, the step change in minimum temperature at Regina Airport, deemed insignificant, appears as a clear régime transition in the rescaled $R$ trace in November 1961. There were no station inspections at either Regina Airport or Midale in 1961, but in November 1962 in response to the question on the Regina Airport inspection form “Is there grass beneath the screen?” someone wrote “no—will be,” implying some recent disturbance to the instrument site, since there was grass in the previous report (1957). Possible instrument relocation is also suggested since the same report notes the length of the psychrometer motor cable to be 600 feet, whereas the previous report stated it was 220 feet long. It seems plausible that the screen may have been moved, perhaps within the same enclosure, but necessitating a longer cable. It is not however possible to determine the precise date of any move but it is likely to be related to construction of a new parking lot and terminal building in 1960/61.

Nineteen sixty-one was also the year in which a nationwide change in observing times occurred at principal stations. Daily maximum and minimum temperature observations changed from the 0000–0000 UTC climatological day to the 0600–0600 UTC day. Biases introduced by this change are discussed by Bootsma (1976). Vincent and Gullett (1999) show that biases resulting from this change were not large enough to affect mean monthly temperatures for provinces west of Manitoba (because of the timing of the observation schedule in relation to the occurrence of the minimum temperature). However, it is interesting that 1961 appears to be the most significant cooling régime transition evident in the rescaled time series of $R$, implying that the observation schedule change may have been
more significant to the homogeneity of the record than analysis of monthly temperature series reveals.

While the 1932 step change documented by G91 is also evident in the rescaled trace, apparently caused by the station move from the City of Regina to the airport site, their method does not identify or consider significant a number of other régime transitions evident in the rescaled time series and/or the original $R$ series. In particular, an inflection around 1939 corresponds to the beginning of significant development at the airport, including expanded runways, new buildings and a control tower. The most significant problem noted in the history files is the heat island problem due to large expanses of surrounding pavement, and gradual encroachment of the site. Although G91 found no statistical evidence of this warming effect in the temperature series, the decreasing trend in the $R$ series between 1964 and 1984 is consistent with the expected thermal influence of extensive paved and built surfaces around the site. The decreasing trend in $R$ ends when the Regina Airport instruments are moved in late 1983, a change that, again, was not identified, or considered significant in G91.

2) EDMONTON MUNICIPAL AIRPORT

G91 also analyze the Edmonton Municipal Airport record (but not that of the International Airport) and find a trend in the maximum temperature series between 1938 and 1983, and step changes in the minimum series in 1942 and 1969. As noted in the Hurst analysis, here there are many régime transitions for EMA in the 1961–99 period. These are apparently caused by station relocations (in 1975, 1984, and 1986) and station automation (in 1994). Hence again the rescaled series of cooling ratios appears able to identify more of the subtle effects that affect cooling régimes at individual sites but are not apparent using conventional homogeneity tests of average temperature data.

3) WOODBEND

Analysis here using the Woodbend, Alberta, record shows that growth of vegetation around a site is likely to be responsible for a sustained increase in the slope of the rescaled $R$ series, and a slight decrease of $R$ itself. G91, on the other hand, conclude after examining many station records that “contrary to popular expectation . . . extensive growth of surrounding vegetation . . .” generally has little or no affect on the resulting data series. It seems that analysis of time series of cooling ratios, including Hurst rescaling, is suited to the identification of such subtle site changes even though they may not be obvious in monthly averages.

6. Conclusions

The preceding results show that analysis of time series of cooling ratios, and in particular, the application of Hurst rescaling to them, is a valuable approach to the detection of subtle, site-specific biases in air temperature observations. This approach constitutes a significant improvement to conventional techniques of homogeneity testing, which often do not identify similar kinds of bias in temperature records.

Estimation of the Hurst exponent indicates whether significant changes to cooling régimes have occurred. If they have examination of both the Hurst-rescaled trace and the cooling ratio time series, in conjunction with station history information, can often determine the cause and timing of the inhomogeneities. Even if, as is often the case, the metadata record is poor the presence of inflections in a time series of rescaled $R$ is important diagnostic evidence of changes in the thermal environment of a station.

Gradual changes in the immediate environment over time, such as vegetation growth, or encroachment by built features such as paths, roads, runways, fences, parking lots, and buildings into the vicinity of the instrument site typically lead to trends in the cooling ratio series. Distinct régime transitions can be caused by seemingly minor instrument relocations (such as from one side of the airport to another, or even within the same instrument enclosure) or due to vegetation clearance. This contradicts the view that only substantial station moves, involving significant changes in elevation and/or exposure are detectable in temperature data (G91). It is not surprising that small station moves, even without changes of elevation or exposure, are capable of introducing inhomogeneities into the record, because there are often several confounding changes occurring at the same time. For example, a station move often coincides with screens being repainted, cleaned, or replaced, new instruments installed, and observers being reinstructed about their practices. Further, it is common for the new instrument site to be without grass for a few years, and there are many indications of muddy conditions around the instruments until grass is both planted and properly maintained. These factors, combined with subtle changes in the immediate surroundings (such as moving away from a parking lot or building), appear to be a significant cause of inhomogeneities in temperature records. As isolated occurrences, activities such as painting, cleaning, or releveling screens or instruments do not frequently cause significant changes to cooling régimes.

This paper is purposely restricted in its scope. Its focus is first to demonstrate both the Hurst rescaling
Hurst rescaling to stations affected by urban heat islands is the subject of a further paper (Oke and Runnalls, unpublished manuscript).

These results reiterate the critical importance of maintaining detailed station history files that can help to identify the cause of temperature and other climatological discontinuities. This is such a basic element of scientific practice and is absolutely essential to enable the full value of climatological observations to be realized. The cost and expertise necessary to archive such simple information is trivial compared to the value forgone due to its absence.

This study suggests that it might be beneficial to re-examine stations that passed previous homogeneity analyses and to consider the implications of the concerns raised by the work here for the large databases of air temperature data that are assumed to be homogeneous and unbiased.

Acknowledgments. This research was funded by the Natural Sciences and Engineering Research Council of Canada through a scholarship to KR and a Discovery Grant to TO. Data and access to the climate station history archives were provided by the Meteorological Service of Canada and we are grateful to them for the help and expertise of Anna Deptuch-Stapf and Lucie Vincent. We thank Prof. Michael Church for bringing Hurst analysis to our attention.

REFERENCES


