Vol. 46: 67–76, 2011
doi: 10.3354/cr00956

CLIMATE RESEARCH
Clim Res

Published online February 8

# Climate fluctuations in time and space

**Igor G. Zurbenko[1,*], Derek D. Cyr[1,2]**

[1]University at Albany, State University of New York, School of Public Health, Rensselaer, New York 12144, USA
[2]Duke University, Center for Applied Genomics and Technology, Durham, North Carolina 27708, USA

ABSTRACT: Satellite imagery of weather was a major breakthrough in the technology of synoptic weather forecasts. Historically, the Earth's climate has experienced dramatic changes, and researchers have found that long-term weather patterns do exist. Such long-term changes in weather patterns have very small variations in total atmospheric variables and cannot be observed directly. Complicated space-time detection of low energy signals is required and can be done only by computationally processing space-time data. In this paper, new statistical methods for processing such data are discussed and applied so as to create images of long-term changes in climate over space. Using the historical Central England temperature (CET) time series, we clearly identified temporal scales of 2 to 5 yr and longer than 13 yr. Using monthly global temperature records obtained from the National Climatic Data Center's Global Historical Climatology Network, a long-term average temperature profile along latitude was identified. Global maps of trends of deviations from the average temperature profile display slowly increasing temperatures over a major part of the world. Maps of 2 to 5 yr scales display deviations similar to those observed during an El Niño event and provide the opportunity for explanation and prediction of weather anomalies in different regions of the world. In this paper, we utilize recent achievements in the technology of processing 3-dimensional data, i.e. in the separation of scales greater than 1 yr in monthly global temperature records. The Kolmogorov-Zurbenko spline filter allows for direct selection of scales in time and space to obtain a smooth outcome without the application of any models.

KEY WORDS: Separation of scales · Kolmogorov-Zurbenko filtration in time and space · Kolmogorov-Zurbenko spline · KZS · El Niño



## 1. INTRODUCTION

Meteorologists often use the term 'scales of motion' to determine whether a physical process is dynamically important in any particular situation. The presence of various scales of motion in meteorological time series can complicate analysis and interpretation of the data. Therefore, it is imperative to be able to separate the time series into synoptic, seasonal and long-term components because the processes occurring at different frequencies are caused by different physical phenomena. In the separation of short-term, seasonal and long-term scales in meteorological variables, Eskridge et al. (1997) showed that by using anomaly techniques based on monthly, seasonal and annual averages, 5% of the energy and 22% of the amplitude of each component can be mistakenly attributed to the others. Rao et al. (1997) described the characteristic space and time scales in time series of ambient ozone data and also discussed the need for a methodology which cleanly separates the various scales of motion embedded in atmospheric data. Separation of subdiurnal scales is detailed by Chapman & Lindzen (1970), Potrzeba & Zurbenko (2008) and Zurbenko & Potrzeba (2009).

The aforementioned papers deal with the separation of scales for the time period of several months to 1 yr. In the present study, we address the separation of scales for exceedance of 1 yr for 100 yr in monthly global temperature records. Monthly records provide more detail than records on synoptic scales and provide sufficient information for precise examination of scales beyond 1 yr. We utilize the Kolmogorov-Zurbenko spline (KZS) spatial filter as the primary tool of analysis. The figures featured in this paper are the direct results of 2 different KZS operations: a 1-dimen-

*Email: igorg.zurbenko@gmail.com

sional (1D) KZS and a 3-dimensional (3D) KZS, the latter of which is the primary filter illustrated. Kolmogorov's idea of filtration was used by North et al. (1995) to separate a signal in time and space from noise using climate data. We identify the parameters of filtration using actual spectral information in regional temperature to identify 2 global signals at different temporal scales.

To provide strong evidence for the existence of these 2 scales, long time series of temperature records are needed. To accomplish this, we use 2 separate data sets. The first is the longest and most accurate temperature time series on record, the Central England temperature (CET) series. We have used this regional time series for a spectral diagnostic of temporal scales and to illustrate separation of scales. This process has revealed 2 scales above seasonal; we have classified these according to long-term global activity and the El Niño phenomenon. These findings are used and extended to the global level by examining them on a second data set containing temperature records of various lengths of time for over 7000 temperature recording stations around the globe via the aforementioned 3D filter.

We are interested in creating spatial deviations of temperature from long-term averages. For this purpose, we have constructed a long-term temperature profile along latitude using 100 yr of monthly global temperature records. Reconstruction of the temporal scales found in the CET data over global data has revealed striking similarities in their pattern in the tropics and Arctic regions.

We also found periodicities that are similar in nature to the irregular periodic oscillations of the El Niño phenomenon in the sea surface temperature (SST) of the tropical Pacific Ocean. Reconstruction of the signals in time and space will help to identify reasons for these oscilllations. Fluctuations in temperature at this assumed El Niño scale are also shown to display some negative correlation with a similar component in precipitation data over the USA, post-El Niño.

## 2. METHODS OF ANALYSIS AND NOTATION

Kolmogorov-Zurbenko (KZ) filters are an ideal methodology for smoothing and/or filtering. Asymptotically, the mean-squared error (MSE) of KZ filters has been shown to be closest to optimal compared with other commonly used filters (Zurbenko 1986, Yang & Zurbenko 2010b). Moreover, Cyr (2010) showed that the asymptotic MSE (AMSE) of the KZS achieves the optimal rate of convergence for nonparametric estimators, according to Stone (1980, 1982). Cyr (2010) demonstrated the advantages of the KZS over other commonly used methods, such as loess, the Gaussian filter and the cubic smoothing spline, in several computationally intensive illustrations. The parameters of the KZS allow for incorporation of a spectral diagnostic into the filter design, as demonstrated in the present study; the other aforementioned techniques have difficulty incorporating spectral language, and they perform poorly in the presence of strong seasonality. Other methods, such as wavelets, are not appropriate for the applications featured here, as the scales discovered in the course of work on this paper are not strictly periodic, but concentrated in certain ranges of frequencies.

For the applications featured in the present study, the 3D KZS filter operates as $k$ iterations of the averages of available gridded information over the $\Delta_x$ span for all longitudes centered at $x$, over the $\Delta_y$ span for all latitudes with center at $y$, and over the $\Delta_t$ span for all times with center at $t$. More formally, the KZS is a function $Z$ having the mathematical expression:

$$\mathrm{KZS}(\Delta,k)[T(x,y,t)] = Z_k[T(x,y,t)] = \frac{1}{|N|}\sum_{i\in\Delta} Z_{k-1}[T(x,y,t)_i] \quad (1)$$

where $T(x, y, t)$ is temperature as a function of $x$, $y$ and time $t$, and $|N|$ denotes the number of $T(x, y, t)_i$ outcome values contained within the 3D smoothing window $\Delta_x \times \Delta_y \times \Delta_t$. The KZS is the spatial extension of the KZ filter (Zurbenko 1986), a low-pass linear filter useful for retaining the slow-moving components of data. The KZ filter, popular for its simple algorithm, operates as $k$ iterations of a moving average of $m$ points. This filter has been used by several investigators to examine the temporal and spatial features embedded in time series of meteorological and air quality data (Eskridge et al.1997, Hogrefe et al. 1998, Zurbenko & Porter 1998, Potter et al. 2000). Capilla (2008) utilized the KZ filter as well as other methods, including wavelets, in a time series analysis of temperature trends in an urban Mediterranean area. For air quality data (see previous references), the KZ filter has been shown to be an extremely effective tool in the analysis of ozone data. Wise & Comrie (2005) have also shown the KZ filter to be an effective tool in the analysis of particulate matter trends as well as an appropriate method of analysis in a semi-arid region of the USA with weaker synoptic controls on air quality compared with the eastern USA, where the KZ filter has already been shown to be effective in the analysis of temperature and ozone (Zurbenko et al. 1995, Rao et al. 1997, Hogrefe et al. 1998, 2000). Because the KZS is based on a moving average, it can perform efficiently in environments with high rates of missing data; this notion is illustrated by Hogrefe et al. (1998) and Cyr (2010). All theoretical foundations of the KZ and KZS filters are provided by Zurbenko (1986), Zurbenko et al. (1995), Cyr & Zur-

benko (2008a) and Yang & Zurbenko (2010b). All figures in the present study were produced using the statistical software R (R Core Development Team 2009, Cyr & Zurbenko 2008b, Close & Zurbenko 2010).

Throughout the present study, the notation KZS($\Delta_x$, $\Delta_y$, $\Delta_t$, $k$) is used to represent the 3D KZS filter. Only the values of the span parameters used in the filter are reported in the displayed order: longitude, latitude, time and number of iterations. If necessary, the span parameters $\Delta_x$, $\Delta_y$ and $\Delta_t$ can be the functions of the center of a 3D cube ($x$, $y$, $t$). For example, to avoid the shrinking effect of distances along longitude when approaching the poles, the manipulation $\Delta_x = \Delta_{x=0}/\cos(y)$ can be taken for any fixed grid; this point is further discussed in Section 4.

When applying a 1D KZS to a single time series, the notation KZS($\Delta$, $k$) is used. Similar notations may be used regardless of the dimensionality of the data as the number of variables in the display determines the dimensionality of the filter.

## 3. PRELIMINARY SPECTRAL ANALYSIS

As a first step in the analysis, we examined the spectrum of the historical CET record. Application of spectral analysis requires long time series and, as a result, we were able to identify 2 low-frequency scales that we will be reconstructing on shorter time series of global data. The CET record contains daily and monthly temperatures of a near triangular region in England. The monthly series was compiled by Manley (1953, 1974) covering the period 1659 to 1973. Parker et al. (1992) updated these data to 1991 while calculating the daily series. Presently, both series are kept up-to-date by the Climate Data Monitoring section of the Met Office Hadley Centre (http://badc.nerc.ac.uk/data/cet/). Furthermore, since 1974, the data have been adjusted by 0.1 to 0.3°C to allow for urban warming. In all, the CET record represents the longest accurate time series of monthly temperature observations (n ~4200) in existence and can act as a stepping-stone for analyses extended to the global level.

The 350-yr length of the time series permits application of high-resolution spectral analysis in the low frequencies. Monthly temperature records are almost certain to contain a very strong seasonal (annual) component; however, this is out of the range of frequencies of interest as we only examined scales exceeding 1 yr in time. This is accounted for in Fig. 1 (described below). Annual and shorter scales were examined in Rao et al. (1997); see also Potrzeba & Zurbenko (2008) and Zurbenko & Potrzeba (2009).

Spectral analysis of the CET data via the Kolmogorov-Zurbenko periodogram (KZP) algorithm (see Yang & Zurbenko 2010a,b) revealed clear evidence of 2 separate components at the following frequency ranges: 2- to 5-yr periods, which we attributed to El Niño oscillations; and temporal scales >13-yr periods, which we attributed to long-term global activity. Fig. 1 displays the resulting KZP spectra from parameters $m$ = 2000, $k$ = 2 and DiRienzo-Zurbenko (DZ) smoothing parameter = 10%, applied within the frequency range of interest (0, 0.05); the huge spike at annual frequency 1/12 is left out of the displayed range and has no influence on the examined frequencies. The advantage of using the KZP in this situation is that it blocks noise leakage from the annual spike to adjacent frequencies. It is worth noting that the regular periodogram always provides a noisy outcome which obscures inferences. Uniform smoothing results in the loss of interesting 'spikes' in the spectrum. We use a 10% smoothing parameter for the DZ algorithm, which allows us to zoom in to spikes above 10% of the total energy of low frequencies (periods >2 yr) and smooth out smaller uncertainties. For further details on the DZ algorithm, see DiRienzo & Zurbenko (1999).

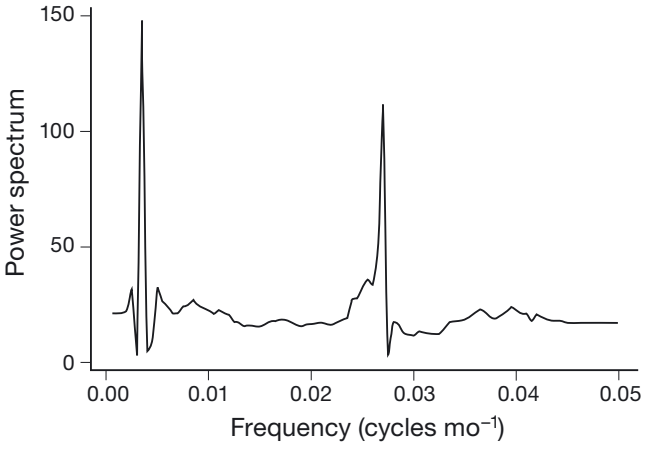


Fig. 1. Spectra of the Central England temperature (CET) data for the monthly time period 1659–2008 resulting from application of the Kolmogorov-Zurbenko Periodogram (KZP) algorithm with parameters $m$ = 2000, $k$ = 2, where $m$ is the length of the window for a regular Fourier transform, and $k$ is the number of iterations of the KZFT algorithm (see details in Yang & Zurbenko 2010b) and DiRienzo-Zurbenko (DZ) smoothing parameter = 10%. Two primary components are identified: long-term scales with frequencies between 0.004 and 0.006, which correspond to a period of >13 yr; and El Niño scales with frequencies between 0.02 and 0.04, which correspond to a period of 2 to 5 yr. The strong, seasonal component is not displayed as it is out of the range of frequencies of interest

The main task illustrated in the present study is the clear separation and restoration of these 2 scales in monthly global temperature records using KZS filtration. We identified the long-term range via KZS

(13 years, 3) as illustrated in Fig. 2. In general, for any arbitrary time series of temperature readings or similar data, this long-term component (LTC) can be written in the following form:

$$\mathrm{LTC}(t) = \mathrm{KZS}(13\mathrm{years}, 3)[T(t)] \quad (2)$$

where $T(t)$ is the monthly mean temperature and $t$ is time, in months. Most of the variation in the CET data is due to seasonality, which was completely annihilated by the chosen filter. The attenuation of the annual frequency in Eq. (2) is extraordinarily small — equal to $10^{-12}$ — so it does not transfer annual seasonality from the data at all. The variance of the long-term component from the CET data is 0.33, which contains 0.45% of the total variance of the CET data. The long-term component extracted by Eq. (2) is highly accurate.

The long-term component provides a very accurate restoration of scales longer than 13 yr, keeping only very small influences of annual and El Niño scales. Calculation of the total remaining variance of the reconstructed long-term signal provides a 95% bracket of confidence in Fig. 2 (dotted line) in the range (–0.2563, 0.2942) towards displayed values.

The same filter operating over annual data provides a near equivalent outcome, with correlation $\rho = 0.99$. Unfortunately, annual data do not provide reliable frequency resolution for long time scales. Periods of 2 to 5 yr are not observable from annual data, as they look like random noise.

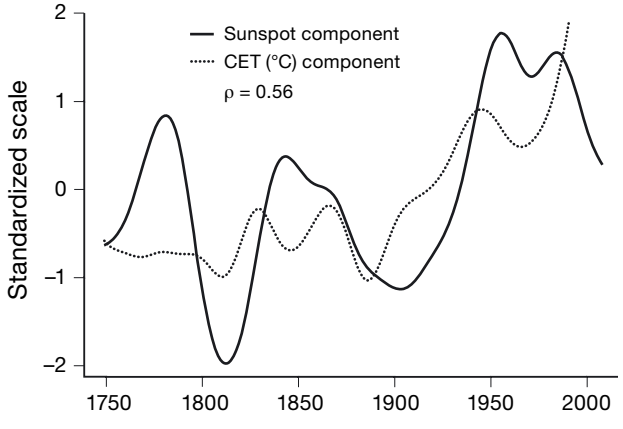


Fig. 2. The long-term component (LTC) extracted from the CET series via a Kolmogorov-Zurbenko spline (KZS) (13 yr, 3) operation exhibits oscillation in scales approximately longer than 13 yr (dotted line). It is evident that the long-term average temperature in this region has increased since 1900. The sunspot LTC over the monthly time period 1749–2008, revealed from a KZS(13 yr, 3) operation is indicated by the black line. Both components are plotted on a standardized scale so that they can be more easily compared. For both components, the mean value was subtracted from each value and then divided by the standard deviation. For the CET component, 1 unit on this scale corresponds to 0.31°C. For the sunspot component, 1 unit corresponds to 15.32 sunspots

LTC over CET data in Fig. 2 can be partially explained by fluctuations in sun activity; it displays a correlation of $\rho = 0.56$ with a similar scale extracted by Eq. (2) over 250 yr of monthly sunspot records obtained from the Solar Influences Data Analysis Center (www.sidc.be/sunspot-data/) (Fig. 2, black curve). The rest of the LTC in temperature may be attributed to other factors such as human activity, a key contributor to global warming.

## 4. DATA PREPARATION

Monthly mean global temperature records were extracted from the National Climatic Data Center's (NCDC) Global Historical Climatology Network (GHCN) version 2 database, which contains historical temperature, precipitation and pressure data for thousands of global land monitoring stations on a monthly time scale; we refer to this database as GHCN-Monthly. An overview of this database is detailed in Peterson & Vose (1997) and Peterson et al. (1998).

With 7280 recording stations, the GHCN-Monthly mean database is more than twice as large as the widely used Jones (1994) 2961-station mean temperature data set (Peterson & Vose 1997).

Data are available for 7280 stations, for which the best spatial coverage is in North America, with the majority of Europe and parts of Australia and eastern Asia having satisfactory coverage as well. The least amount of coverage occurs in the polar regions of the globe, especially for the first half of the 20th century. All stations have at least 10 yr of recorded data.

As the spatial distribution of temperature recording stations is uneven, the data are gridded to create a uniform space over the globe. At the same time, the data are adjusted to sea level to compensate for the various changes in elevation throughout the globe, using the dry adiabatic lapse rate (DALR).

In many meteorological models, a grid is used to represent data for an area. For example, the data from a single temperature recording station can be considered to be representative of the conditions for a 50-mile (~80 km) square block around that station. By using data from several stations, a grid can be constructed to cover a large land area with a reasonable level of accuracy. For each month in the GHCN-Monthly database, we create a 5° latitude × 5° longitude grid to respectably cover the globe, following from the work of Peterson & Vose (1997). This corresponds to a grid consisting of 36 latitude rows by 72 longitude columns, for a total of 2592 grid cells for each month. Where data are available, the mean temperature for each station is placed in the respective grid cell. For grid cells that contain data for multiple recording stations, the mean value of those readings is taken and is assigned to the respective cell.

## 5. SEPARATION OF SCALES IN GLOBAL TEMPERATURE RECORDS

To obtain support for our findings with the CET series, we began examination of the gridded GHCN-Monthly temperature data. Despite the fact that most of the cells within the spatial grids are empty, KZ filtration has the advantage of being able to handle missing data. The KZS treats all grids equally and without exclusion, so by filtration of a specific scale, it provides an outcome for $T(x, y, t)$ over all time $t$ and all spatial grids defined by longitude $x$ and latitude $y$. This is equivalent to obtaining an outcome of a specific scale distributed uniformly over all grids.

The first step in the creation of uniform $T(x, y, t)$ information is the application of KZS(15°, 15°, 1 yr, 3) over the 5° × 5° gridded GHCN-Monthly data. This operation fills each grid on a monthly basis, with equal representation in each grid. The time scale of 1 yr was chosen in order to smooth out seasonal scales, leaving only longer scales to address. The function $T(x, y, t)$ represents all scales in global temperature in the range of 2 to 100 yr and spatial distances over 1500 miles; shorter scales have been eliminated by the averaging process. Also, the function contains both El Niño and global scales, evenly presented in time and space. A similar regional construction in regards to smoothing ozone and temperature over the USA is given in Zurbenko et al. (1995). For the purposes of the present study, $T(x, y, t)$ acts as a starting point for all further analyses and represents uniform information in space and time for temporal scales longer than 2 yr.

## 6. GLOBAL TROPICAL AND ARCTIC LONG-TERM COMPONENTS

The primary global regions of study are the tropics (in this case, 25° N to 25° S) and all high north latitudes (Arctic, 60 to 90° N) because, especially in the Northern Hemisphere, temperature recording stations have become more prevalent from year to year, thus providing more accurate results than in southern latitude regions.

For each global region, a 100 yr average monthly temperature time series was created by averaging all available data across longitude. Using Eq. (2), the LTCs were extracted and are shown in Fig. 3. This figure clearly displays a long scale wave similar to what was received from the single time series of the CET data in Fig. 2, over the same time period (CET curve in Fig. 3). The LTC for the tropical and Arctic regions is smoother than the CET LTC because these regions use more data, which contributes to the high confidence of these outcome curves. The correlation between these 2

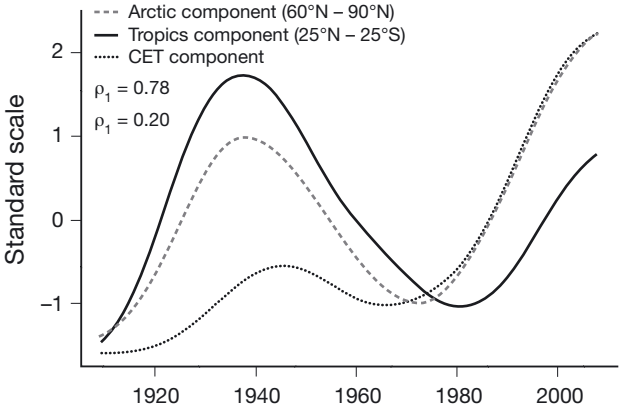


Fig. 3. Overlay of the CET long-term component (LTC) with Arctic and tropical LTC. Both components are plotted on a standardized scale so that they can be more easily compared. For the tropics component, 1 unit on this scale corresponds to 0.09°C. For the Arctic component, 1 unit corresponds to 0.31°C

components is strong, $\rho = 0.71$. Further, the correlation between the CET and Arctic LTCs is also strong, $\rho_1 = 0.78$; however, the correlation between the CET and tropical LTCs is much lower, $\rho_2 = 0.20$. This large difference may be due to the fact that England (the source of the CET series) is geographically closer to the Arctic than to the tropics, thus providing more similar temperature measurements. Nevertheless, the nonparametric trend in all 3 curves is extremely similar, providing strong evidence that the LTC found in the CET data has a global scale.

## 7. LONG-TERM LATITUDINAL TEMPERATURE DISTRIBUTION AND VISUALIZATION OF 4D DATA

A map, $T(x, y)$, of the 100-yr average of $T(x, y, t)$, which depicts the long-term average temperature distribution over the globe is given in Fig. 4. Fig. 5 depicts the average of $T(x, y, t)$ over the available record of time $t$ and all longitudes $x$ as an outcome function $T(y)$ of latitude $y$. We also approximate this long-term average parametrically (°C) by the function:

$$T(y) \sim \{[a\cos^2(y) + b] - 32\} \times 5/9 \quad (3)$$

where $a = 63.35$, $b = 12.94$ and $R^2 = 0.99$; that is, 99% of the variation in $T(y)$ is accounted for by Eq. (3). The largest gradient along latitude of this approximation occurs at latitude $y = 45°$ and is equal to approximately $^5/_9$°C per 1° of latitude. For latitudes close to the equator and poles, the gradient vanishes to zero. The accuracy in the approximation by $\cos^2(y)$ is high enough such that Eq. (3) may be considered a law of average temperature distribution along latitudes.

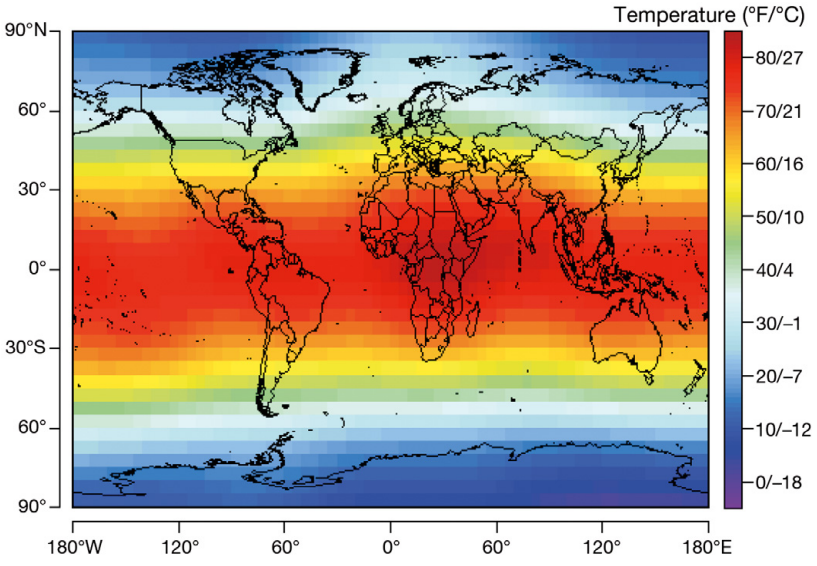


Fig. 4. Global map of smoothed 100-yr monthly mean temperature distribution $T(x, y)$, where $x$ represents longitude and $y$ latitude. This was produced via application of KZS to the 5° × 5° grid containing the 100 yr mean temperature in each grid cell over the monthly time period January 1909 through December 2008, where data were available. This grid contained data in 922 (35.6%) of the total 2592 grid cells. Here, KZS used $\Delta$ along longitudes proportional to $1/\cos(y)$. Specifically, for 30° N to 30° S latitude, $\Delta = 15°$; for 30° to 60° in both hemispheres, $\Delta = 25°$; for 60° to 90° in both hemispheres, $\Delta = 55°$. Celsius readings have been rounded to the nearest degree

A rough 'physical sense' may support this squared cosine law. At the equator, any unit area $S$ receives the maximum amount of (shortwave) radiation because the sun is directly overhead (on the zenith). As latitude $y$ increases away from the equator, $S$ turns away from a straight angle and receives radiation proportional to $\cos(y)$. At the same time, at latitude $y$, the sunlight is traveling in the atmosphere proportionally to $1/\cos(y)$ and the intensity of sun radiation is again proportional to the distance of travel of light, which yields $1/[1/\cos(y)] = \cos(y)$. Thus, we have $\cos^2(y)$ as a factor of attention to the equatorial radiation at latitude $y$, which we can also use to approximate temperature profiles over latitude.

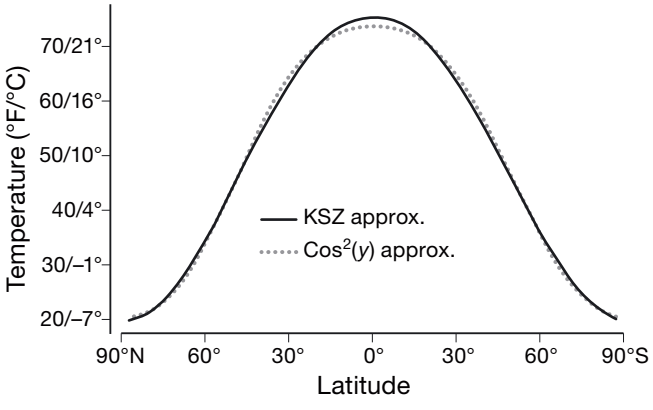


Fig. 5. Distribution of mean temperature over all latitudes, January 1909 through December 2008. To produce the KZS approximation, data from Fig. 4 have been averaged across longitude, yielding the black line, $T(y)$. The grey dotted line is a parametric approximation (°C) according to the expression $T(y) \sim [(63.35\cos^2(y) + 12.94) - 32] \times 5/9$ ($R^2 = 0.99$), which can be considered a law of average temperature distribution along the latitudes. Celsius readings have been rounded to the nearest degree

Eq. (3) crosses the freezing point of temperature, 0°C (32°F), approximately at latitude 57°N. We consider this as the theoretical permafrost line, which, in reality, may fluctuate slightly depending on regional long-term temperature deviations.

The most convenient way to visualize 4-dimensional (4D) data containing geographic components is through a 'movie' (see Henry et al. 1997, available at ftp://ftp.dec.state.ny.us/dar/air_research/htdocs, tmpg folder). Values of $T(x, y, t)$ of 4D set $\{T, x, y, t\}$ are plotted on a 2-dimensional map of latitude and longitude $\{(x, y)\}$, with the values of temperature $T$ scaled in color. Monthly time $t$ is used as an indicator of different maps shown consecutively and actual time is displayed at the top of each image when playing a series of images such as can be observed in Henry et al. (1997). We subtract out the long-term profile given in Eq. (3) in order for small fluctuations to become visible in a restrictive color scale.

## 8. THE GLOBAL LONG-TERM COMPONENT

The global LTC can be displayed as a 'movie' of deviations of KZS(15°, 15°, 13 yr, 3)[$T(x, y, t)$] from the approximation given by Eq. (3). The deviations are relatively small, which makes such a movie possible via colored maps shown consecutively. In general, this 'movie' reveals extraordinary stability for the last 50 yr, while also displaying leaks of energy from the tropics to higher latitudes. Noticeable changes over time are apparent in observation of these maps over 10-yr increments. Fig. 6 shows maps of the global LTC for June 1967 and June 2007. The negative deviations seen in the June 1967 map have shrunk considerably over time and, as is visible in the June 2007 map, where a majority of the globe is covered by positive deviations. From these maps, it is evident that global warming has different rates in different areas.

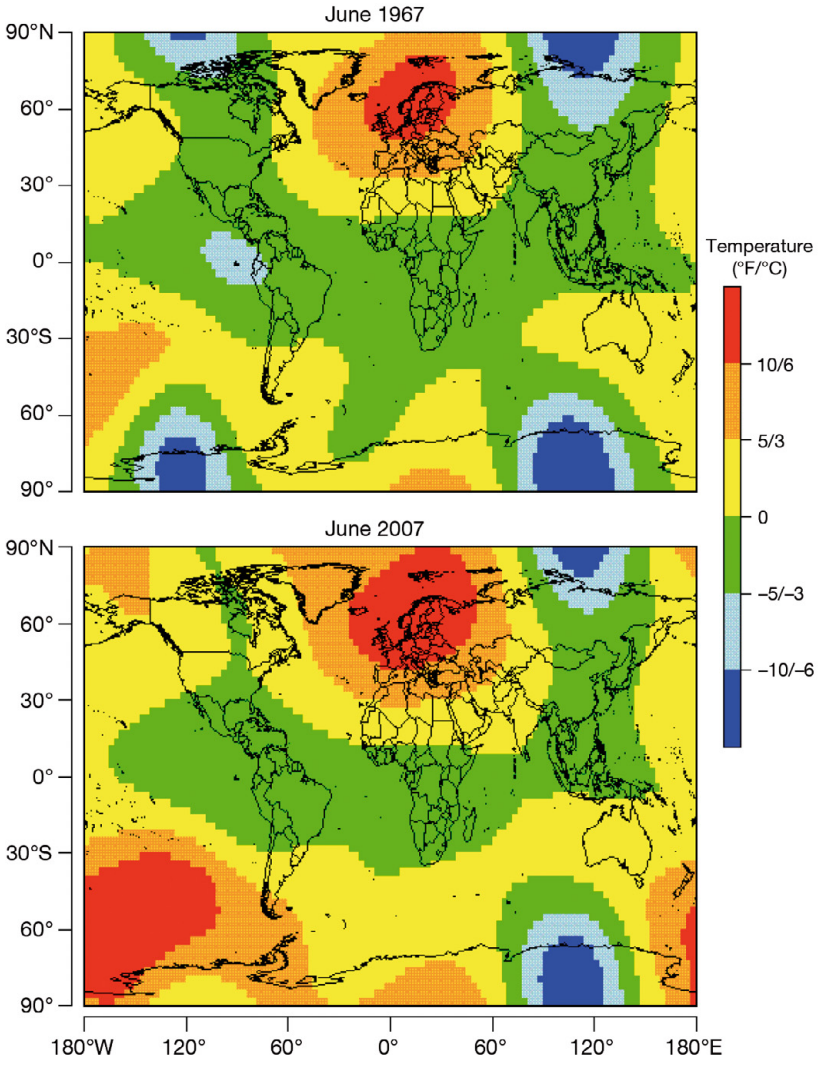


Fig. 6. Deviation in the long-term component. Maps reveal very slowly changing deviations from cosine law around the entire globe. Visualization of the entire collection of monthly images as a series of images can provide a possible forecast for one decade ahead in time. The large positive deviation in the Antarctic is likely due to the spatial sparseness of temperature recording stations throughout the 100 yr time period. As a result, grid cells in this region use available data from the nearest source, which in this case is the tropics. For the entire 100 yr time period, the Northern Hemisphere has approximately 2 to 3 times more available data than the Southern Hemisphere. Celsius readings have been rounded to the nearest degree

## 9. EL NIÑO-SCALE OSCILLATIONS

The El Niño-scale component can be displayed as a 'movie' of deviations $T(x, y, t)$ – KZS(15°, 15°, 13 yr, 3) $[T(x, y, t)]$. The El Niño-scale is essentially a time and spatial signal with irregular periods of 2 to 7 yr, which can be conveniently restored by the aforementioned 3D KZS operation. Uneven distribution of energy over the planet will inevitably cause some deviation from the stable curve approximated by the squared cosine function. Some areas will become warmer than usual, causing drier conditions and, potentially, drought. Kim et al. (2003) discuss a nonparametric approach to the return periods of droughts as an effect of El Niño cycles. Furthermore, evaporated moist air will try to move to cooler areas, creating more cloud cover for those areas and restricting the sun radiation as an energy supply. As a result, warmer areas will receive more energy for a particular period of time, causing drier conditions, whereas cooler areas, experiencing moist conditions and restricted energy supply, become cooler. This yields oscillations in time and space, which can be presented as a series of images.

Penland & Matrosova (2006) utilized spatial filtration to separate tropical SSTs into 3 components — a non-normal El Niño signal, a global tropical trend and the background SST anomaly field — while accounting for a high amount of variation in the El Niño signal. Atmospheric temperature yields a more challenging problem of filtration, where the El Niño scale counts for only single percentage points of total variation. The high resolution provided by the KZS filter can overcome this problem.

For simplicity, consider the USA, where a dense collection of data permits decreasing of the $\Delta$ parameter to yield finer details in the outcome maps. In this case, a 0.5° latitude × 0.5° longitude grid is defined over the 48 contiguous states. The time period illustrated here is 1978–1997, a period during which multiple strong El Niño events occurred. More than 50% of the grids contain no information and are classified as missing. Spatial parameters of filtration are selected in order to avoid local weather fluctuations of scales less than 500 miles (~800 km). This corresponds to a KZS operation with $\Delta_x$ = 2.5°, $\Delta_y$ = 2.5°. Further, we specify $\Delta_t$ = 1 yr and $k$ = 3 iterations. The simplistic nature of the KZS filter is essential here; it has been able to detect time oscillations in space with 2 to 5 yr periods. These oscillations can be classified as El Niño-scale oscillations and may be predictable up to 1 yr in advance.

The El Niño scale can be attributed to the re-distribution of energy over the surface of the Earth. It has scales of 2 to 5 yr in time and approximately 500–2000 miles (~805 to 3220 km) in space, and displays fluctuations around the global LTC. Some similarities

between this scale in the atmosphere and water temperature in the oceans have been observed, but more details should be further examined. Statistically, the global scale and El Niño scale are uncorrelated and should be investigated separately, as we did in the present study. Possible interactions between these 2 scales cannot be investigated until each one has been fully reconstructed in time and space.

Fig. 7a illustrates this outcome for June 1993, a point in time during the middle of the 'Great Flood of 1993' in the Mississippi River basin, and just after the conclusion of the 1991–1993 El Niño event in the tropical Pacific Ocean. A strong negative temperature anomaly is clearly visible. This negative build-up was noticeable and predictable from KZS images starting in late 1992. A movie constructed from these images is available from Henry et al. (1997).

For this same region, we should also be able to observe positive deviations in precipitation measurements. Using data obtained from the US Historical Climatology Network (USHCN) monthly data version 2 database, monthly mean precipitation measurements from 1218 recording stations were gridded in the same fashion as above over the same time period, and the same KZS operation was applied. The 20 yr monthly average precipitation value was subtracted from each grid cell so that deviations from long-term local averages are observable. Fig. 7b is the outcome of this KZS operation for June 1993. Visually, there is evidence of high negative correlation, primarily in the Mississippi River basin region, as indicated by the strong negative anomalies in temperature and the positive anomalies in precipitation.

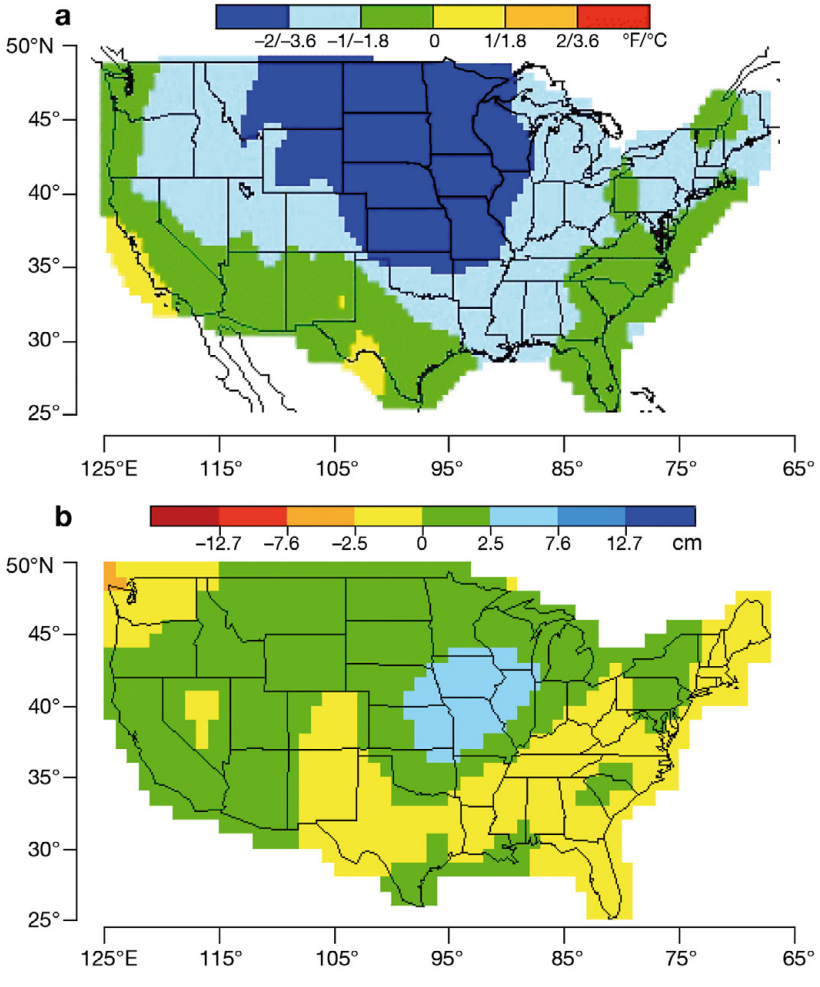


Fig. 7. (a) Deviations in smoothed temperature from mean values (1978–1997) over the USA during the 1993 summer season, post 1991–1993 El Niño event. A strong negative temperature anomaly in the Mississippi River basin is clearly visible, reflective of the 'Great Flood of 1993,' which occurred between April and October of that year. This negative build-up was noticeable and predictable from KZS images starting in 1992. (b) Deviations in smoothed precipitation from mean values (1978–1997) for June 1993. Throughout the Mississippi River basin, positive deviations in precipitation levels are visible at the same time as the negative temperature deviations

## 10. LIMITATIONS

The GHCN-Monthly database provides a very uneven distribution of available information over the globe. For 5° × 5° grids, the rate of missing information exceeds 35% over the 100 yr monthly time period, which hardly permits the construction of more desirable smaller grids; this is only possible for areas with a dense network of temperature monitors. The map of the USA illustrating El Niño oscillations provides such an opportunity with 0.5° × 0.5° grids creating a 'high-definition' movie. For the entire globe, such a construction appears impossible due to lack of information. However, smaller grids, such as the 0.5° × 0.5° grids, capture spatial fluctuations in the El Niño-scale much better.

Spherical coordinates (longitude $x$, latitude $y$) on the globe have discontinuities at the date separation line and at the poles. Grid size shrinks to 0 as $y$ approaches the poles. Nevertheless, KZS is capable of providing evenly distributed results per unit area, even with such complications. By taking the total width parameter $\Delta_x$ along $x$ proportional to $1/\cos(y)$, an even treatment of all weighted areas on the globe will result. This solution has been applied in Fig. 4. Spherical distances may be used

in order to avoid discontinuities over the date line. To avoid these discontinuities, KZS may read longitudinal coordinates as −179 = 181. After such treatment, outcomes will become continuous over the date line.

## 11. CONCLUSIONS

The Heisenberg Uncertainty Principle states that the product of the time interval of data times the interval of accuracy of frequency determination remains constant (Folland & Sitaram 1997). In other words, there is a limit to frequency resolution. To achieve better resolution, the sample size must be increased. In this case, 100 yr of data is not enough to prove the existence of the 2 temporal scales illustrated in the present study; however, the 350 yr length of the CET series is enough for such a purpose. This allows for the application of a decomposition filter over 100 yr worth of data for restoration of those scales.

The long period covered by the CET monthly time series records permits investigation of its spectra in frequency ranges of 2 to 200 yr periods. High-resolution spectral analysis revealed 2 strongly elevated frequency ranges of approximately 2 to 5 yr and greater than 13 yr. Using KZ filtration, ranges of long-term periods and trends exceeding 2 to 5 yr can be reproduced with great accuracy. Outcomes are nonparametric and can be used for any further analysis as a separate variable. This allows forecasts in time of up to half of the considered time scale and the examination of associations between different atmospheric variables within those scales in space.

One hundred years of monthly global temperature records from the GHCN support these findings. Filtration using a 3D KZS enabled reproduction of all scales above 2 yr with continuous coverage in space $T(x, y, t)$, which is approximately adjusted to unit area on the surface of the Earth. Such a reconstruction of temperature in time and space is relevant to the energy in the low-level atmosphere in unit area.

The LTCs averaged over the entire gridded tropics region (25°N to 25° S) and the Arctic (60 to 90° N) revealed a very smooth oscillatory pattern over the last 100 yr and are very similar to the LTC extracted from the CET data over the same time period. These findings have led to the conclusion that the LTC in temperature has a global scale.

The average of $T(x, y, t)$ over 100 yr and all gridded longitudes revealed a distinctive, smooth function over latitude $y$ that was shown to be closely approximated parametrically by $\cos^2(y)$. The accuracy in the approximation given by Eq. (3) is high enough to be considered a law of long-term average temperature distribution along latitude.

A fine-scale movie of temperature over the USA displays slow oscillations of temperature as deviations of 20 yr local averages; some areas appear cooler, whereas others appear warmer, with periodicity of 2 to 5 yr, corresponding to the El Niño scale. Such deviations can be predictable up to 2 yr in advance. Long-term prediction of weather is possible from these movies; however, such forecasts address large-scale patterns of weather rather than separate events.

The LTC and El Niño scales are very well hidden in the total variation in atmospheric data. In most cases, seasonality alone is responsible for more than 90% of the total variation. KZ filtration is specially designed for extremely fine resolution between different frequencies and can solve the problem of reconstruction of fine components in low-frequency ranges. The total value of these components is low, but their energy influences are high enough to contribute to climate change.

A main feature of KZ filtration is that it provides very strong suppression of noise contained in a given data set or time series (see Yang and Zurbenko 2010a,b). Large simulation studies conducted by Cyr & Zurbenko (2008a) and Potrzeba & Zurbenko (2008) have shown a very high accuracy of reconstructed signals from noisy environments. The level of noise shown in Figs. 6 and 7 does not exceed single percentage points of the reconstructed values of the signals. The width of the filters applied in these maps is in the range of 500 miles (~805 km) (Fig. 6) and 1500 miles (~2415 km) (Fig. 7), so outcomes on longer distances display high correlation, although they are obtained from completely different areal data. Nevertheless, these maps show coherent wave structures of scales of several thousand miles. These structures are not artifacts; they are long-term and El Niño-scale fluctuations in time and space. KZ filtration is a nonparametric technique, so no models have been applied to obtain the observed fluctuations. They were simply reconstructed from measured data as low-frequency components. Within a period of 1 yr from the last data observation, they can be forecasted in time and space.

*Acknowledgements.* The authors thank Dr. Robert F. Henry from the New York State Department of Environmental Conservation for his valuable insights that helped to make this paper come together.

LITERATURE CITED

Capilla C (2008) Time series analysis and identification of trends in a Mediterranean urban area. Global Planet Change 63:275–281

Chapman S, Lindzen RS (1970) Atmospheric tides. Gordon and Breach, New York

Close B, Zurbenko IG (2010) kza: Kolmogorov-Zurbenko adaptive filters. R package version 1.02. Available at http://cran.r-project.org

Cyr D (2010) A spline kernel based smoothing algorithm: a comparison of methods and a spatiotemporal application to global climate fluctuations. PhD thesis, University at Albany, State University of New York

Cyr D, Zurbenko IG (2008a) A spatial spline algorithm and an application to climate waves over the United States. Proc Joint Statistical Meetings 2008, Denver, CO, Aug 3–7, 2008. American Statistical Association, Alexandria, VA

Cyr D, Zurbenko IG (2008b) kzs: Kolmogorov-Zurbenko spatial smoothing and applications. R package version 1.4. Available at http://cran.r-project.org.

DiRienzo G, Zurbenko IG (1999) Semi-adaptive nonparametric spectral estimation. J Comput Graph Stat 8:41–59

Eskridge R, Ku J, Rao ST, Porter PS, Zurbenko IG (1997) Separating different scales of motion in time series of meteorological variables. Bull Am Meteorol Soc 78:1473–1483

Folland G, Sitaram A (1997) The uncertainty principle: a mathematical survey. J Fourier Anal Appl 3:207–238

Henry R, Sanapati S, Zurbenko IG (1997) 4-dimensional images of El Niño oscillations in the air temperature over the USA. Available at ftp://ftp.dec.state.ny.us/dar/air_research/htdocs. (See tmpg folder)

Hogrefe C, Rao ST, Zurbenko IG (1998) Detecting trends and biases in time series of ozonesonde data. Atmos Environ 32:2569–2586

Hogrefe C, Rao ST, Zurbenko IG, Porter PS (2000) Interpreting information in ozone observations and model predictions relevant to regulatory policies in the Eastern United States. Bull Am Meteorol Soc 81:2083–2106

Jones PD (1994) Hemispheric surface air temperature variations: a reanalysis and an update to 1993. J Clim 7: 1794–1802

Kim TW, Yoo C, Valdés JB (2003) Nonparametric appoach for estimating effects of ENSO of return periods of droughts. KSCE J Civ Eng 7:629–636

Manley G (1953) The mean temperature of Central England, 1698 to 1952. Q J R Meteorol Soc 79:242–261

Manley G (1974) Central England temperatures: monthly means 1659 to 1973. Q J R Meteorol Soc 100:389–405

North GR, Kim K, Shen SSP, Hardin JW (1995) Detection of forced climate signals. I. Filter theory. J Clim 8:401–408

Parker DE, Legg TP, Folland CK (1992) A new daily Central England temperature series, 1772-1991. Int J Climatol 12: 317–342

Penland C, Matrosova L (2006) Studies of El Niño and interdecadal variability in tropical sea surface temperatures using a nonnormal filter. J Clim 19:5796–5815

Peterson TC, Vose RS (1997) An overview of the Global Historical Climatology Network temperature database. Bull Am Meteorol Soc 78:2837–2849

Peterson TC, Vose RS, Schmoyer R, Razuvaev V (1998) Global Historical Climatology Network (GHCN) quality control of monthly temperature data. Int J Climatol 18:1169–1179

Potrzeba A, Zurbenko IG (2008) Algorithm of periodic signal reconstruction with applications to tidal waves in atmosphere. Proc Joint Statistical Meetings 2008, Denver, CO, Americal Statistical Association, Alexandria, VA

Potter BW, Raman S, Niyogi DS (2000) Applying the KZ filter for studying North Carolina temperature and precipitation patterns associated with ENSO. Proc Joint 15th Conf on Probability and Statistics in the Atmospheric Sciences/12th Conf on Applied Climatology, May 8–11 2000, Ashville, NC. American Meteorological Society, Boston, MA, p J16–J19

R Core Development Team (2009) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rao ST, Zurbenko IG, Neagu R, Porter PS, Ku JY, Henry RF (1997) Space and time scales in ambient ozone data. Bull Am Meteorol Soc 78:2153–2166

Stone CJ (1980) Optimal rates of convergence for nonparametric estimators. Ann Stat 8:1348–1360

Stone CJ (1982) Optimal global rates of covergence for nonparametric regression. Ann Stat 10:1040–1053

Wise EK, Comrie AC (2005) Extending the KZ filter: application to ozone, particulate matter, and meteorological trends. J Air Waste Manag Assoc 55:1208–1216

Yang W, Zurbenko IG (2010a) Nonstationarity. Comp Stat 2: 107–115

Yang W, Zurbenko IG (2010b) Kolmogorov-Zurbenko filters. Comp Stat 2:340–351

Zurbenko IG (1986) The spectral analysis of time series. Elsevier, Amsterdam

Zurbenko IG, Porter PS (1998) Construction of high-resolution wavelets. Signal Process 65:315–327

Zurbenko IG, Potrzeba A (2009) Tidal waves in the atmosphere and their effects. Acta Geophysica 58:356–373

Zurbenko IG, Rao ST, Henry RF (1995) Mapping ozone in the eastern United States. Environ Manag 1:24–30

*Editorial responsibility: Bryson Bates, Wembley, Australia*

*Submitted: March 9, 2010; Accepted: October 18, 2010*
*Proofs received from author(s): January 17, 2011*