Graphical display and statistical modeling of temperature changes in tropical and subtropical zones

Cherdchai Me–ead1,2 and Nittaya McNeil1,2*

1 Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Mueang, Pattani, 94000 Thailand.

2 Centre of Excellence in Mathematics, Commission on Higher Education (CHE), Ministry of Education, Ratchathewi, Bangkok, 10400 Thailand.

Received: 6 November 2015; Accepted: 21 March 2016

Abstract

Climate change, particularly rising temperature, is one of the important environmental problem facing the world today. This study aims to identify trends and patterns of temperature change in tropical and subtropical zones using a statistical model. Data were obtained from the climate research unit from 1973 to 2008, comprising 252 regions of 10° by 10° grid-boxes between latitudes 35° north and south. The data were filtered with a second order autoregressive process to remove autocorrelations between temperature lags. Factor analysis was used to classify monthly average temperature anomalies into larger regions by taking into account the correlation between adjoining regions. Simple linear regression models were then fitted to the data within these larger regions. The result showed that the temperatures in these 15 larger regions have increased the most (by at least 0.065°C per decade) in the North Atlantic Ocean and the central and the north of Africa. Lower increases (0.045–0.064°C per decade) occurred in Southeast Asia, the Indian Ocean and the North Pacific Ocean.

Keywords: linear regression model, autocorrelation, factor analysis, climate change, tropical zones

1. Introduction

Global warming is an important issue in the world today. Rising average surface temperatures is one of many indicators of global warming, causing sea levels to rise, melting of snow and icebergs (the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and Bureau of Meteorology, 2007), droughts in the Amazon, inundation of coastal cities, large-scale ecosystem transformations (Moutinho et al., 2005) and extreme fluctuations in the El Niño Southern Oscillation index (Hughes, 2003).

Since the 1950s, the tropics have become a domain of permanently increasing interest among meteorologists (Sitnikov, 2009) mainly because the weather in the tropics is dominated by convective storms. These develop mainly along the Intertropical Convergence Zone. The tropical and subtropical zone is surrounded by the three largest oceans; the Atlantic, Pacific and Indian oceans; and large continental regions (Pidwirny, 2011). The different temperature variability on the earth’s surface in the tropical and subtropical zones has been reported in several studies. For example, sea surface temperatures in most tropical cyclone formation regions have increased during the past decades (Knutson et al., 2010). The ocean surface temperatures around Australia have increased 0.7°C since 1910-1929, with south-western and south-eastern waters warming fastest (Poloczanska et al., 2009). The average temperature increase of 0.26°C occurred in the upper ocean during the 1957-1998 period at around 24.5°N and average cooling of -0.15°C was found in the period of 1988-2005 (Vélez-Belchí, 2010). The temperature from 1951-1975...
increased sharply by about 3.10°C in the Pacific Ocean (Hartmann and Wendler, 2004). Several studies investigated possible trends and patterns of the earth surface temperature changes using different statistical methods such as the Mann-Whitney U test (James and Washington, 2013), ordinary least squares (OLS) regression, Durbin–Watson statistic (Collins, 2011), Mann-Kendall test (Collins, 2011; Tshiala et al., 2011; Klein Tank et al., 2006), linear regression analysis, Pearson correlation analysis (Mote, 2003; Houghton and Tourre, 1991), principal components analysis (Houghton and Tourre, 1992), multiple linear regression analyses (Lean and Rind, 2009), linear spline functions, multivariate linear regression models and factor analysis (McNeil and Chooprateep, 2014), and autoregressive moving average: ARMA (Hughes et al., 2007).

Investigating the temporal patterns of temperature change can indicate the zones with increased temperature and possibly predict the periods of likely global temperature change (global warming). Information on earth surface temperature change and the interpretations of these trends by time period and geographic region is essential for temperature forecasts, environmental policy and many other purposes. Therefore, the objective of the study is to investigate the trends and patterns of temperature change in tropical and subtropical zones using appropriate statistical models.

2. Materials and Methods

2.1 Data management and study areas

Temperature data of the 36-year period from 1973 to 2008 were obtained from the Climate Research Unit (CRU, 2009). The CRU provides monthly temperature averages for 5° by 5° latitude-longitude grid-boxes on the earth’s surface based on data collected from weather stations, ships, and, more recently, satellites. Temperature data were composed of 432 monthly average temperature anomalies, defined as increases over the monthly averages for the period 1961-1990. Data were combined into 252 regions of 10° by 10° latitude-longitude grid-boxes which were designed in the shape of igloo bricks. These 252 regions were located on the earth’s surface between latitudes 35 north and south, which cover the tropical zone. Thus, there were 252 regions for the analysis as shown in Figure 1. There were missing data in some regions for several months. Shading (lightest, intermediate and darkest) was used to denote the percentage of missing data. Of the 252 regions, 224 regions (white) contained complete data, 24 regions (lightest shade) < 10% missing data, and three regions (intermediate shade) had more than 10% missing data. One region (darkest shade), in central South America, had no data and was omitted from further analysis.

2.2 Statistical methods

A simple linear regression was used to fit the monthly average temperature anomalies for each of the 251 regions. The linear model is:

\[ y_{it} = b_{0i} + b_{1i}d_t, \]

for \( i = 1, 2, 3, ..., 251 \) and \( t = 3, 4, 5, ..., 432 \)

where \( y_{it} \) denotes the monthly average temperature anomaly in region \( i \), month \( t \); \( d_t \) denotes the elapsed time centered in the middle of the period; that is \( d_t = (t-n/2)/120 \) for a period of 432 months (36 years). Thus \( b_{0i} \) is the monthly average temperature anomaly in the grid-box over the period; and \( b_{1i} \) is a corresponding rate of increase in temperature anomaly per decade.

Second order autoregressive models were used to obtain the autocorrelations in the residuals from the fitted linear models. The monthly average temperature anomalies \( (y_{it}) \) were filtered to remove autocorrelation at lags 1 and 2 months in each region. This model takes the form:

\[ z_{it} = a_{1i}z_{i,t-1} + a_{2i}z_{i,t-2} \]

for \( i = 1, 2, 3, ..., 251 \) and \( t = 3, 4, 5, ..., 432 \)

where \( z_{it} = y_{it} - (b_{0i} + b_{1i}d_t) \) is the residuals in region \( i \) and at month \( t \).

Factor analysis (Mardia et al., 1980) was used to classify the filtered monthly average temperature anomalies of the 251 regions into larger regions by taking account of correlation between adjoining regions. The factor analysis model with \( p \) factors takes the form:

Figure 1. Earth’s surface between latitudes 35° north and south, different shading denotes the percentage of missing data.
Data analysis and graphical displays were carried out using the R language and environment (R Development Core Team, 2009).

3. Results and Discussion

The coefficients of slope ($b_i$) were obtained from the simple linear regression for each of the 251 regions. Figure 2 shows the 95% confidence intervals for monthly average temperature anomalies ordered by longitude with each latitude band. Most of the 95% confidence intervals were higher than zero indicating significant temperature increase (238 regions). In twelve regions the 95% confidence interval was less than zero indicating a significant temperature decrease. The percent of missing data in each region is shown by square points on the grey horizontal line. The temperature change in the northern hemisphere (a, b and c) had different trends and patterns from the southern hemisphere (d, e and f) and the equatorial region (g). There were similar patterns in each latitude band.

![Figure 2. 95% confidence intervals for monthly average temperature anomalies change per decade ordered by longitude within each latitude band. The black rectangle shows region studied in each range group of latitude.](image)
in the northern hemisphere as well as in southern hemisphere. Temperature changes in Africa of the northern hemisphere were high as compared to southern hemisphere. A similar finding was also observed in the Atlantic Ocean where high temperature changes were evident. However, in the Pacific Ocean, temperature changes were higher in the southern Pacific than in the northern Pacific. At the equator, temperature change in the Pacific Ocean was similar to temperature change of Pacific Ocean of the northern hemisphere and temperature change in Atlantic Ocean was similar to temperature change of Atlantic Ocean of the southern hemisphere.

Figure 3 shows an autocorrelation plot of residuals from the linear regression model for four of the 251 regions. They were selected from different regions to be representative of the Pacific Ocean, Atlantic Ocean, Indian Ocean and South America, respectively. The residuals were assumed to be stationary. The autocorrelation function plot showed that most of the correlations were outside the 95% confidence interval in the early lags. The remaining regions also had most of the correlations outside the 95% confidence interval.

The autocorrelation was removed by using a second order autoregressive process for all 251 regions. Then coefficient \( a_1 = 0.4921 \) and \( a_2 = 0.1028 \) was obtained from all region and the average was taken. The coefficients \( a_1 \) and \( a_2 \) were used as filtered temperature. Figure 4 shows the plot of autocorrelation of residuals from linear regression of filtered temperature for the four regions shown in Figure 3. It was found that autocorrelation was within the 95% confidence interval. Thus, the autocorrelation were removed as shown in this Figure. Similarly, autocorrelation was also removed in the remaining 247 regions.

Spatial correlation was considered between temperature change in 251 regions. So, factor analysis was used to classify the filtered temperature of 251 regions into 15 larger regions by taking account of spatial correlation among adjacent regions. Figure 5 plots the 15 larger regions after using factor analysis and coded from A to O. Seventy five regions had high uniqueness and could not be classified into larger regions.

Figure 6 shows the 95% confidence intervals of temperature change per decade did not include zero, indicating significant increased temperatures.

The 15 larger regions were re-classified into four levels by quartile ranking into four equal groups according to the lower bound of the 95% CI of its temperature change per decade: there were less than 0.060°C, 0.061°C-0.110°C, 0.111°C-0.160°C, and more than 0.161°C temperature change. Four colours (green, yellow, orange and red) were used to indicate the four different temperature increased as shown in Figure 7. Blue colours represent high uniqueness regions from the factor analysis.

Four classifications of temperature change were obtained. First, the maximum temperature increased by at least 0.161°C per decade in the north Atlantic Ocean, the central and northern regions of Africa, and south west Asia (red larger regions C, O and E). The results are consistent with the studies in the same region by Collins (2011), Maiyza and Kemel (2006), Smadi (2006) and Deser et al. (2010).

Figure 3. Autocorrelation plot of the residuals for four of the 251 regions. The dotted line represents the 95% confidence interval for a zero correlation.

Figure 4. Autocorrelation plots of the filtered residuals for four of the 251 regions. The dotted line represents the 95% confidence interval for a zero correlation.
Second, the temperature increases (0.111-0.160°C per decade) covered the south of Asia, the Arabian sea, the bay of Bengal, the Indian ocean, south east Asia, and the north Pacific ocean (orange larger regions G, A, H and K), suggesting high temperature increase. The results of increased temperatures are similar to the temperature trend in Bangladesh for the period 1948-2007 (Islam, 2009), 1961-2008 (Shahid et al., 2012); India for the period 1944-1994 (Arora et al., 2005); and, tropical Pacific during 1950-2010 (L’Heureux et al., 2013). Third, the increase rate of temperature (0.061-0.110°C per decade) was seen in the southern part of the north American continent, the gulf of Mexico, the Caribbean sea, the part of the north Atlantic ocean, the south Atlantic ocean, the central and western of Africa, and part of the south pacific, which covered the coral sea (yellow larger region I, F, L and J). These results are similar to the studies by Insaf et al. (2013), Collins (2011) and Kruger and Shongwe (2004). The lowest increase in temperature (below 0.060°C per decade) was detected in the north and south Pacific ocean, central and south America and Australia (green larger region D, B, M and J). The results of rising temperature were similar to the studies of Vincent et al. (2005) and Deser et al. (2010).

4. Conclusions

The statistical model used to investigate temperature changes in this study involved removing autocorrelations of monthly average temperature anomaly data with the second order autoregressive process and accounting for the spatial correlation by using factor analysis to classify temperatures of 251 regions into 15 larger regions. A simple linear regression model then was used to fit temperature data in each larger region.

The results revealed that the monthly temperature in tropical and subtropical zones had increased in all larger regions ranging from 0.012 to 0.627°C during 1973 to 2008. The different patterns were related to position of latitudinal variations of the climate. However, in this study the missing
data were managed by combining grid-boxes. This could cause loss of information because it averages out the estimation of the temperature. Further study could use different statistical methods for handling missing data and explore the temperature trend of each region. The temperature change of land and sea could also be investigated separately.

Acknowledgements

We are grateful to Emeritus Prof. Don McNeil, Prince of Songkla University, Thailand, for providing valuable advice and support. We would like to thank the Centre of Excellence in Mathematics, Commission on Higher Education, Thailand, for providing support for this research.

References


