INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD • INDIA

Are the temperature of Indian cities increasing?: Some insights using Change point analysis with Functional Data

Arnab Kumar Laha

Poonam Rathi

W.P. No.2017-08-03

August 2017

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380 015
INDIA

Are the temperature of Indian cities increasing?: Some insights using Change point analysis with Functional Data

Arnab Kumar Laha
Poonam Rathi

Production and Quantitative Methods Area
Indian Institute of Management Ahmedabad
arnab@iima.ac.in
poonamr@iima.ac.in

Abstract

In recent years there has been considerable concern expressed worldwide regarding increase in temperature popularly called the global warming problem. In this paper we examine monthly temperature data of nine Indian cities for the period 1961 to 2013. We introduce a new Gaussian process based method for change point detection with functional data and use it to investigate the existence of change point for the temperature data series of nine Indian cities. It is found that there has been a rise in the average temperature for eight of the nine cities during this period. The magnitude of warming is found not to be uniform but vary across cities located in different parts of India. The cities located in hilly areas is seen to have warmed more than those located in the plains. The estimated change points for the eight cities are not identical but most of them are in the period 1994 - 2001. The findings suggest that immediate policy measures are required to ensure that no further warming happens in these cities.

Keywords: Global warming, Climate change, Gaussian process, Powered Exponential Covariance Function, Functional Principal Component Analysis, Generalized Likelihood Ratio Test

1 Introduction

In recent years, a lot of attention of the world's leaders both in developing and developed countries has been drawn towards "climate change". The Intergovernmental Panel on Climate Change (IPCC) uses this term to indicate change in climate (i.e. the planet's weather patterns or average temperatures) over time, whether due to natural variability or as a result of human activity. A major concern which has received wide coverage in the popular press across the world is the phenomenon of global warming which can be described as an increase in average surface temperature compared to the past, possibly due to human activity. While there is no consensus amongst climate scientists regarding the exact magnitude of this increase and its causes, there are a large number of them who believe that the global warming is a real phenomenon. Climate scientists from the IPCC estimate that average global temperatures could increase between 1.1°C and 6.4°C by the year 2100 compared to the 1980-1999 levels, IPCC (2007). The effects of global warming may include rising sea levels due to the melting of the polar ice caps leading to nearly complete submersion of low-lying countries (such as Maldives), as well as an increase in occurrence and severity of extreme weather events such as storms, drought etc.

In recent years we have a resurgence of interest in reducing global warming and mitigating the impact of climate change. The Kyoto Protocol, which is an international treaty that commits countries to reduce greenhouse gas emissions was adopted in December 1997 and entered into force in February 2005. The 2010 United Nations Climate Change Conference (UNCCC) held in Cancun, Mexico adopted an agreement that called for the countries to create a large Green Climate Fund, and a Climate Technology Centre and network. The recently held 2015 UNCCC in Paris, France resulted in an agreement that committed all the countries to work towards reduction of their greenhouse gas emissions in amounts that meets the aim of restricting global warming to below 2°C.

In India, climate change is thought to be impacting the natural ecosystems. It is expected that in the long run substantial adverse effects mainly on, agriculture on which 58 per cent of the Indian population still depends for livelihood, water availability for cultivation of crops and drinking due to melting of the Himalayan glaciers which are the source of India's major rivers, and loss of habitation and food security due to sea-level rise resulting in inundation of large tracts of fertile coastal lands with sea water.

India being the seventh largest country in the world with north-south distance more than 3200 km and east-west distance more than 2900 km it is expected that the effect of global warming would vary across different regions of the country. In this paper we consider nine cities from different parts of India and examine if the yearly average temperature has risen during the period 1961-2013 in these cities. The findings indicate that global warming may become a cause of concern to a large section of the population of India. We further investigated whether the temperature increase varies across seasons and regions. We find that the temperature increase is more in the winter season indicating the winters have become milder in many parts of the country. Moreover, we find that the temperature increase is more for high altitude regions than that of plains. Similar studies have addressed temperature changes in other countries like Finland [Mikkonen et al. (2015)].

The remainder of the article is organized as follows. In section 2, we provide a background on stochastic process, functional data and change point problem and in section 3, we provide a brief review of Gaussian process with powered exponential covariance function that is relevant for this paper. In section 4, we propose a method for detecting the presence of change point using Generalised Likelihood Ratio test. In section 5, we analyze the monthly temperature data for the period 1961-2013 for nine Indian cities to detect the presence of a change point. In section 6, we explore the relation of amount of change with geographical location, seasons as well as altitude followed by some concluding remarks in section 7.

2 Background

A stochastic process indexed by a set T is a collection $X = \{X(t)\}_{t \in T}$ of measurable maps from a probability space (Ω, F, P) with values in a measurable space (E, ξ). $X(t)$ is called a random element which is a generalization of the concept of a random variable (where $(E, \xi) = (\mathbb{R}, B)$). The set $X(\omega) = \{X(t, \omega) : t \in T\}$ is called a sample path of the process. In the context of this paper we take E to be the space of all continuous square integrable real valued functions. We consider the temperature at a place over a year (the temperature curve) as an element of E. Thus we visualize that for each year we have an element of E which is drawn at random according to the probability distribution P.

In functional data analysis (FDA) the data is represented in the form of curves unlike that in the conventional univariate or multivariate data where the observations are either scalars or vectors. There has been a substantial progress in the statistical analysis of functional data over the last three decades. FDA has been applied in a variety of scientific fields like medicine [Erbas et al. (2009)], biology [Müller et al. (2009)], environment [Gao and Niemeier (2008)], ecology [Ikeda et al. (2008)] and many other areas. Ramsay and Dalzell (1991) analyses daily temperature and precipitation levels at 35 Canadian weather-stations over a year using a functional data approach. Meiring (2007) uses FDA to study altitude variation in the ozone partial pressure profiles, which also includes nonlinear time trends. Applications of FDA has greatly increased in the past few years because of rapid advancement of data gathering technologies such as sensors, and increase in processing power of computers. The books by Hsing and Eubank (2015), Horváth and Kokoszka (2012), Ramsay et al. (2009), Ramsay and Silverman (2002), Ramsay and Silverman (2005) offer a broad perspective of the available methods and case studies in functional data.

We consider functional observations $X_i(t)$, $t \in T, i = 1, \ldots, n$ defined over [a,b]. It is assumed that X_i are independent and identically distributed as X, drawn from $L^2([a,b])$. Suppose that the mean function be $\mu = E(X) \in L^2([a,b])$. Further, assume that the covariance function K is a continuous function on $[a,b] \times [a,b]$. Then, there exists a sequence of continuous eigenfunctions ϕ_n and a decreasing sequence of corresponding non-negative eigenvalues λ_n such that

$$\int_a^b K(s,t)\phi_n(s)ds = \lambda_n \phi_n(t), \qquad \int_a^b \phi_n(s)\phi_m(s)ds = \delta_{nm}$$

Also, each functional observation can be decomposed as $X(t) = \mu(t) + \sum_{n=0}^{\infty} \eta_n \phi_n(t)$ where (η_n) is a sequence of real zero-mean random variables such that $E(\eta_n \eta_m) = \lambda_n \delta_{nm}$. Moreover, $K(s,t) = \sum_{n=0}^{\infty} \lambda_n \phi_n(s)\phi_n(t)$; $s,t \in [a,b]$; where the series converges uniformly and absolutely on (a,b) [Bosq (2000),(p.25)].

The change point problem for the mean function with functional observations was studied in Berkes et al. (2009) using a non-parametric approach. Suppose $X_i(t), t \in T, i = 1, 2, \ldots, n$ be a temporally ordered sequence of independent functional observations defined over a compact set T with mean functions $\mu_1, \ldots, \mu_n$ respectively. We are interested in testing the following hypothesis:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_n \text{ versus the alternative } H_1 : \text{not } H_0$$

The test statistic $S_{n,d}$ given in Berkes et al. (2009) rejects H_0 if the value of the test statistic is greater than the tabled values given in table 1 of Berkes et al. (2009) where

$$S_{n,d} = \frac{1}{n^2} \sum_{\ell=1}^{d} \frac{1}{\hat{\lambda}_\ell} \sum_{k=1}^{n} \left(\sum_{i=1}^{k} \hat{\eta}_{i,\ell} - \frac{k}{n} \sum_{i=1}^{n} \hat{\eta}_{i,\ell} \right)^2 \tag{1}$$

where $\hat{\eta}_{i,\ell}$ are the estimated functional scores expressed as

$$\hat{\eta}_{i,\ell} = \int_T [X_i(t) - \overline{X}_n(t)]\hat{\phi}_n(t)dt$$

where the quantity $\hat{\phi}_n$ is estimated as explained earlier and d is chosen to explain around 85% of variance while applying principal component analysis. Aue et al. (2009) provide a method for estimating the change point based on maximizing a quadratic form.

3 Gaussian Process

In the context of FDA, the outcome for each experiment can be considered as realization of a sample path of the underlying stochastic process. A convenient assumption is to take the underlying process to be a Gaussian Process. The class of Gaussian processes is one of the most widely used families of stochastic processes for modeling dependent data observed over time (see for e.g. Müller and Yang (2010), Shi and Choi (2011)). A Gaussian process $\{X(t), t \in T\}$, indexed by a set T (in this paper we take T to be the set of non-negative real numbers), is a stochastic process, in which any finite linear combination of random variables $X(t)$, have a joint multivariate normal distribution. Equivalently, $\{X(t), t \in T\}$ is a Gaussian process, if for any choice of distinct values $t_1, \dots, t_k \in T$, the random vector $X = (X(t_1), \dots, X(t_k))^T$ has a multivariate normal distribution with mean vector $\mu = E(X) = (E(X(t_1)), \dots, E(X(t_k)))^T$ and covariance matrix $\Sigma = (Cov(X(t_i), X(t_j)))_{i,j=1,\dots,k} = (\sigma_{ij})_{i,j=1,\dots,k}$. The mean and covariance functions of a Gaussian process are given by

$$\mu(t) = E(X(t)) \qquad \text{and}$$

$$\Sigma(s,t) = Cov(X(s), X(t)) = E(X(s) - E(X(s)))(X(t) - E(X(t)))$$

respectively. A Gaussian process is completely specified by its mean function and covariance function. Among the many desirable properties associated with the Gaussian process is the Karhunen-Loeve (KL) expansion. The KL-expansion of a centered Gaussian process $\{X(t), t \in T\}$ can be represented as [Wahba (1990)(p.5)]

$$X(t) = \sum_{k=1}^{\infty} \xi_k \phi_k(t)$$

where $\xi_1, \xi_2 \dots$ are independent, Gaussian random variables with

$$E\xi_k = 0, \quad E\xi_k^2 = \lambda_r$$

and

$$\xi_k = \int_T X(s)\phi_k(s)ds, \ \Sigma(s,t) = \sum_{k=1}^{\infty} \lambda_k \phi_k(s)\phi_k(t) \text{ and } \int_T\int_T \Sigma^2(s,t)\, ds\, dt < \infty$$

where λ_k is a decreasing sequence of non-negative eigenvalues and ϕ_k is a corresponding sequence of continuous eigenfunctions.
In this paper, we restrict ourselves to the powered exponential covariance function which is defined as

$$\Sigma(s,t) = v_0 \exp(-w|s-t|^\gamma), \qquad 0 < \gamma \leq 2,\ v_0 > 0,\ w > 0 \tag{2}$$

We will denote a Gaussian process with powered exponential covariance function as $GP_{PE}(\mu, v_0, w, \gamma)$. In applications, we will assume that all observations are independent and each observation is a sample path of a $GP_{PE}(\mu, v_0, w, \gamma)$. The parameters (μ, v_0, w, γ) are usually unknown and they are to be estimated from the given data.

3.1 MLE of mean and covariance parameters of $GP_{PE}(\mu, v_0, w, \gamma)$

Assume that the curves $X_1, \ldots, X_n$ is a random sample from $GP_{PE}(\mu, v_0, w, \gamma)$ and each of these curves are observed at same time points $t_1, \ldots, t_k$. The log likelihood of the data is given by:

$$l_0 = l(\tilde{\mu}, v_0, w, \gamma) = \sum_{i=1}^{n} \left(\frac{-k}{2} ln2\pi - \frac{1}{2} ln|\Sigma| - \frac{1}{2} (\tilde{X}_i - \tilde{\mu})' \Sigma^{-1} (\tilde{X}_i - \tilde{\mu}) \right) \tag{3}$$

where $\tilde{X}_i = (X_i(t_1), \ldots, X_i(t_k))$, $\tilde{\mu} = (\mu(t_1), \ldots, \mu(t_k)), \Sigma$ is the $k \times k$ matrix whose $(g,h)^{th}$ element is $v_0 \exp(-w(t_g - t_h)^\gamma)$. Differentiating equation (3) with respect to $\tilde{\mu}$ and equating to zero yields $\hat{\mu}(t_l) = \frac{\sum_{i=1}^{n} X_i(t_l)}{n}$, $1 \leq 1 \leq k$. Again, differentiating equation (3) with respect to v_0 and equating to zero yields $\hat{v}_0 = \frac{\sum_{i=1}^{n} (\tilde{X}_i - \hat{\mu})' T^{-1} (\tilde{X}_i - \hat{\mu})}{nk}$ where $\Sigma = v_0 T$ and T is the $k \times k$ matrix whose (g, h)$^{\text{th}}$ element is $\exp(-w|t_g - t_h|^\gamma)$. In case w and γ are assumed known then $\hat{v}_0$ is the MLE of v_0.
Writing $\Sigma = \left[\sigma_{ij}\right]_{k\times\ k}$ where $\sigma_{ij} = v_0 \exp(-wc_{ij}^\gamma)$ where $c_{ij} = |t_i - t_j|$, $l_0 = l_0(\mu, \sigma_{11}, \sigma_{12}, \ldots, \sigma_{1k}, \sigma_{22}, \sigma_{23}, \ldots, \sigma_{2k}, \ldots, \sigma_{kk})$ and $\sigma_{ij} = \sigma_{ij}(v_0, w, \gamma)$ if $i \neq j$ else $\sigma_{ij} = v_0$. Since

$$\frac{\partial l_0}{\partial w} = \sum_{i,j=1}^{k} \frac{\partial l_0}{\partial \sigma_{ij}} \frac{\partial \sigma_{ij}}{\partial w}, \qquad \frac{\partial l_0}{\partial \gamma} = \sum_{i,j=1}^{k} \frac{\partial l_0}{\partial \sigma_{ij}} \frac{\partial \sigma_{ij}}{\partial \gamma} \tag{4}$$

$$\frac{\partial \sigma_{ij}}{\partial w} = -c_{ij}^\gamma \sigma_{i,j} \quad \text{and} \quad \frac{\partial \sigma_{ij}}{\partial \gamma} = -wc_{ij}^\gamma \log(c_{ij}) \sigma_{i,j} \tag{5}$$

by using Smith (1978),

$$\frac{\partial l_0}{\partial \sigma_{ij}} = \frac{1}{2} \text{tr} \left\{ \left[-n\Sigma^{-1} + \Sigma^{-1} \left(\sum_{m=1}^{n} \left(\tilde{X}_m - \tilde{\mu} \right) \left(\tilde{X}_m - \tilde{\mu} \right)' \right) \Sigma^{-1} \right] \left[\frac{\partial \Sigma}{\partial \sigma_{ij}} \right] \right\} \tag{6}$$

we can obtain the estimates of w and γ by substituting values from equation (5) and equation (6) in equation (4) and then equating it to zero and solving for w and γ. As the analytical solutions are not tractable, numerical techniques can be used to obtain these estimates. Alternatively, one may attempt to maximize the log-likelihood function directly using a global optimization algorithm. Among the several modern numerical approaches to global optimization of functions the Differential Evolution (DE), Storn and Price (1997) approach is one of the most promising. In the R software, the package DEoptim provides the functionality of optimizing a given function using the DE algorithm. We use the DEoptim algorithm for obtaining the MLEs of the parameters in this paper motivated by Mullen (2014) where it is reported that the performance of the DEoptim algorithm is good across a variety of benchmark problems.

4 Change point detection

Let us assume that we have a time-ordered sequence of n independent functional observations $X_1, \ldots, X_n$. We assume that these observations come from a Gaussian process with powered exponential covariance function. In actual applications with real data often it is found that the functional observations X_i are not observed at all-time points but only at a few discrete points $0 \leq t_{i1} < t_{i2} < \ < t_{ik_i}$ and in general, it is possible that the points t_{ij} and the constants k_i depend on i. However for the data-sets analyzed in this paper, k_is are all equal (= k say) and $t_{ij} = t_{kj}$ for all i and k, and for each j, $1 \leq j \leq k$. We are interested to determine if there exists a point r, $1 \leq r \leq n$ such that

$$X_1, \ldots, X_r \sim GP_{PE}(\mu_0, v_0, w_0, \gamma_0)$$

$$X_{r+1}, \ldots, X_n \sim GP_{PE}(\mu_1, v_1, w_1, \gamma_1)$$

We adopt a generalized likelihood ratio test (GLRT) methodology for this problem.

Let $\Omega = \{(\boldsymbol{\mu_0},\ \boldsymbol{\mu_1},\ v_0,\ v_1,\ w_0,\ w_1,\ \gamma_0,\ \gamma_1, r) : \boldsymbol{\mu_0},\ \boldsymbol{\mu_1} \in \mathbb{R}^k,\ v_0 > 0,\ v_1 > 0,\ w_0 > 0,\ w_1 > 0,\ 0 < \gamma_0 \leq 2,\ 0 < \gamma_1 \leq 2,\ 1 \leq r \leq n\}$, $\Omega_0 = \{(\boldsymbol{\mu_0},\ \boldsymbol{\mu_1} = \boldsymbol{\mu_0},\ v_0,\ v_1 = v_0,\ w_0,\ w_1 = w_0,\ \gamma_0, \gamma_1 = \gamma_0) : \boldsymbol{\mu_0},\ \boldsymbol{\mu_1} \in \mathbb{R}^k, v_0 > 0,\ w_0 > 0,\ 0 < \gamma_0 \leq 2,\ r = n\}$ and $\Omega_1 = \Omega - \Omega_0$. It may be noted that $\Omega_0 \cap \Omega_1 = \varnothing$ and $\Omega_0 \cup \Omega_1 = \Omega$.

Further, let $L^* = \sup_{(\theta \in \Omega)} L(\theta)$, $L_0^* = \sup_{(\theta \in \Omega_0)} L(\theta)$ and $L_1^* = \sup_{(\theta \in \Omega_1)} L(\theta)$ where L denotes the likelihood function. We wish to test $H_0 : \theta \in \Omega_0$ versus $H_1 : \theta \in \Omega_1$.

The generalized likelihood ratio(GLR) test statistic is defined to be $\lambda = \frac{L_0^*}{L^*}$ and the GLRT rejects H_0 for small values of λ. An equivalent test can be based on $\ln \lambda^* = l_1^* - l_0^*$ where $l_0^* = \ln L_0^*$ and $l_1^* = \ln L_1^*$ which rejects H_0 for large values of $\ln \lambda^*$. We use $\ln \lambda^*$ as the GLR test statistic in this paper.

4.1 Change point problem for mean when covariance function is unchanged

We first consider the simplified case where no change is envisaged for the covariance function i.e. it is assumed that $v_0 = v_1, w_0 = w_1$, and $\gamma_0 = \gamma_1$. Suppose we have observed X_1 at time points $t_{11}, t_{12}, \ldots, t_{1k_1}$; X_2 at time points $t_{21}, t_{22}, \ldots, t_{2k_2}$; $\ldots, X_n$ at time points $t_{n1}, t_{n2}, \ldots, t_{nk_n}$, i.e. possibly at different time points. Then the log likelihood under H_0 is given by

$$l_0 = l(\mu_0, v_0, w_0, \gamma_0) = \sum_{i=1}^{n} \left(\frac{-k_i}{2} ln2\pi - \frac{1}{2} ln|\Sigma_{0i}| - \frac{1}{2} (\tilde{X}_i - \tilde{\mu}_{0i})' \Sigma_{0i}^{-1} (\tilde{X}_i - \tilde{\mu}_{0i}) \right) \quad (7)$$

where $\tilde{X}_i = (X_i(t_{i1}), \ldots, X_i(t_{ik_i}))$, $\tilde{\mu}_{0i} = (\mu_0(t_{i1}), \ldots, \mu_0(t_{ik_i}))$, Σ_{0i} is the $k_i \times k_i$ matrix whose (g, h)$^{\text{th}}$ element is $v_0 \exp(-w_0|t_{ig} - t_{ih}|^{\gamma_0})$. Under H_{1r} where H_{1r} is the alternative hypothesis with change at r, the log likelihood l_{1r} is given by

$$\begin{aligned} l_{1r} = l(\mu_0, v_0, w_0, \gamma_0, \mu_1) = &\sum_{i=1}^{r} \left(\frac{-k_i}{2} ln2\pi - \frac{1}{2} ln|\Sigma_{0i}| - \frac{1}{2} (\tilde{X}_i - \tilde{\mu}_{0i})' \Sigma_{0i}^{-1} (\tilde{X}_i - \tilde{\mu}_{0i}) \right) + \\ &\sum_{i=r+1}^{n} \left(\frac{-k_i}{2} ln2\pi - \frac{1}{2} ln|\Sigma_{0i}| - \frac{1}{2} (\tilde{X}_i - \tilde{\mu}_{1i})' \Sigma_{0i}^{-1} (\tilde{X}_i - \tilde{\mu}_{1i}) \right) \end{aligned} \quad (8)$$

where $1 \leq r \leq n-1$, $\tilde{X}_i = (X_i(t_{i1}), \ldots, X_i(t_{ik_i}))$, $\tilde{\mu}_{0i} = (\mu_0(t_{i1}), \ldots, \mu_0(t_{ik_i}))$, $\tilde{\mu}_{1i} = (\mu_1(t_{i1}), \ldots, \mu_1(t_{ik_i}))$, Σ_{0i} is the $k_i \times k_i$ matrix whose $(g, h)^{\text{th}}$ element is $v_0 \exp(-w_0|t_{ig} - t_{ih}|^{\gamma_0})$. Then the GLR test statistic can be expressed as $\ln \lambda^* = \max_{1 \leq r \leq n}(l_{1r} - l_0)$, where $l_{1n} \equiv l_0$. The position of the change point is estimated by $\hat{r}$ where $\hat{r}$ is the value of r for which $\ln \lambda_r^* = l_{1r} - l_0$ attains its maximum. The null distribution of $\ln \lambda^*$ is not analytically tractable. In case $(\mu_0, v_0, w_0, \gamma_0)$ are known then the cut-off points of the test can be obtained using simulation. If μ_0 is known but v_0, w_0, γ_0 are unknown, then an approach similar to Silvapulle (1996) can be taken to overcome the problem of nuisance parameters. In case all the nuisance parameters under H_0 are unknown, an adaptive approach can be pursued to obtain approximate cut-off points as discussed later in this paper.

Note that if each $X_1, \ldots, X_n$ are defined at same time points $t_1, \ldots, t_k$,then the mean of the function estimated from equation (8) is $\hat{\mu}_0(t_l) = \frac{\sum_{i=1}^{r} X_i(t_l)}{r}$, and $\hat{\mu}_1(t_l) = \frac{\sum_{i=r+1}^{n} X_i(t_l)}{n-r}$, $1 \le l \le k$. The v_0 of the function estimated from equation (8) is

$$\hat{v}_0 = \frac{\sum_{i=1}^{r}(\tilde{X}_i - \hat{\mu}_{0i})' T^{-1} (\tilde{X}_i - \hat{\mu}_{0i}) \ + \ \sum_{i=r+1}^{n}(\tilde{X}_i - \hat{\mu}_{1i})' T^{-1} (\tilde{X}_i - \hat{\mu}_{1i})}{nk}$$

where $\Sigma = v_0 T$ where T is the $k \times k$ matrix whose (g, h)$^{\text{th}}$ element is $\exp(-w_0|t_{ig} - t_{ih}|^{\gamma_0})$.

It may be noted that for this simplified situation an analogue of this problem in the mutivariate normal set-up but without any parametrization of the covariance matrix has been discussed in Srivastava and Worsley (1986) (see also Zamba and Hawkins (2006)).

4.2 Change point problem for mean when covariance function may also have changed

We now consider the general case of the change point problem where we want to detect a possible change in the mean and / or covariance function. We assume that in case both mean function and covariance function have changed, the changes have occurred at the same time point. Suppose we have observed X_1 at time points $t_{11}, t_{12}, \ldots, t_{1k_1}$; X_2 at time points $t_{21}, t_{22}, \ldots, t_{2k_2}, \ldots, X_n$ at possibly time points $t_{n1}, t_{n2}, \ldots, t_{nk_n}$, i.e. at different time points. Then the log likelihood under H_0 is given by

$$l_0 = l(\mu_0, v_0, w_0, \gamma_0) = \sum_{i=1}^{n} \left(\frac{-k_i}{2} ln 2\pi - \frac{1}{2} ln|\Sigma_{1i}| - \frac{1}{2}(\tilde{X}_i - \tilde{\mu}_{0i})' \Sigma_{1i}^{-1} (\tilde{X}_i - \tilde{\mu}_{0i}) \right) \quad (9)$$

where $\tilde{X}_i = (X_i(t_{i1}), \ldots, X_i(t_{ik_i}))$, $\tilde{\mu}_{0i} = (\mu_0(t_{i1}), \ldots, \mu_0(t_{ik_i}))$, Σ_{0i} is the $k_i \times k_i$ matrix whose (g,h)th element is $v_0 \exp(-w_0(t_{ig} - t_{ih})^{\gamma_0})$.
Under H_{1r} where H_{1r} is the alternate hypothesis with change at r, the log likelihood l_{1r} is given by

$$
\begin{aligned}
l_{1r} &= l(\mu_0, v_0, w_0, \gamma_0, \mu_1, v_1, w_1, \gamma_1) \\
&= \sum_{i=1}^{r} \left(\frac{-k_i}{2} ln2\pi - \frac{1}{2} ln|\Sigma_{0i}| - \frac{1}{2} (\tilde{X}_i - \tilde{\mu}_{0i})' \Sigma_{0i}^{-1} (\tilde{X}_i - \tilde{\mu}_{0i}) \right) + \\
& \sum_{i=r+1}^{n} \left(\frac{-k_i}{2} ln2\pi - \frac{1}{2} ln|\Sigma_{1i}| - \frac{1}{2} (\tilde{X}_i - \tilde{\mu}_{1i})' \Sigma_{1i}^{-1} (\tilde{X}_i - \tilde{\mu}_{1i}) \right)
\end{aligned} \tag{10}
$$

where $\tilde{X}_i = (X_i(t_{i1}), \ldots, X_i(t_{ik_i}))$, $\tilde{\mu}_{0i} = (\mu_0(t_{i1}), \ldots, \mu_0(t_{ik_i}))$, $\tilde{\mu}_{1i} = (\mu_1(t_{i1}), \ldots,$ $\mu_1(t_{ik_i}))$,Σ_{0i} is the $k_i \times\ k_i$ matrix whose (g,h)th element is $v_0 \exp(-w_0(t_{ig} - t_{ih})^{\gamma_0})$, Σ_{1i} is the $k_i \times k_i$ matrix whose (g,h)$^{\text{th}}$ element is $v_1 \exp(-w_1(t_{ig} - t_{ih})^{\gamma_1})$.

Again note that if $X_1, \ldots, X_n$ are observed at same time points $t_1, \ldots, t_k$, the mean of the function estimated from equation (10) is $\hat{\mu}_0(t_l) = \frac{\sum_{i=1}^{r} X_i(t_l)}{r}$, and $\hat{\mu}_1(t_l) = \frac{\sum_{i=r+1}^{n} X_i(t_l)}{n-r}$, $1 \le l \le k$. The v_0 is estimated from equation (8) is $\hat{v}_0 = \frac{\sum_{i=1}^{r} (\tilde{X}_i - \hat{\mu}_0)' T_0^{-1} (\tilde{X}_i - \hat{\mu}_0)}{rk}$ where $\Sigma_0 = v_0 T_0$ where T_0 is the $k \times k$ matrix whose (g,h)$^{\text{th}}$ element is $\exp(-w_0 |t_g - t_h|^{\gamma_0})$. The v_1 is estimated from equation (8) is $\hat{v}_1 = \frac{\sum_{i=r+1}^{n} (\tilde{X}_i - \hat{\mu}_1)' T_1^{-1} (\tilde{X}_i - \hat{\mu}_1)}{(n-r)k}$ where $\Sigma = v_1 T_1$ where T_1 is the $k \times k$ matrix whose (g,h)$^{\text{th}}$ element is $\exp(-w_1 |t_g - t_h|^{\gamma_1})$.

As discussed earlier, the GLR test statistic can be expressed as $\ln \lambda^* = \max_{1 \le r \le n}($ $l_{1r} - l_0)$ where $l_{1n} \equiv l_0$. The position of the change point is estimated by $\hat{r}$ where $\hat{r}$ is the value of r for which $\ln \lambda_r^* = l_{1r} - l_0$ attains its maximum. The null distribution of $\ln \lambda^*$ is not analytically tractable. In case $(\mu_0, v_0, w_0, \gamma_0)$ are known then the cut-off points of the test can be obtained using simulation. In case all the nuisance parameters under H_0 are unknown an adaptive approach can be pursued to obtain approximate cut-off points as discussed later in this paper.

5 Warming of Indian Cities

In this section we study the monthly temperature data of nine Indian cities for the period 1961-2013 and examine the possible existence of a change point using the technique discussed in section 4.2. This data is obtained from the Indian Meteorological department (IMD), Ministry of Earth Sciences, Government of India. The nine cities under consideration are Ahmedabad, Bengaluru, Imphal, Jaipur, Kolkata, Port Blair, Pune, Srinagar and Trivandrum. Table 1 shows the latitude, longitude, altitude and location of these cities. These cities are located in different parts of the country and are at differing altitudes.

City	Latitude	Longitude	Altitude(m)	Region
Ahmedabad	23.02	72.57	53	Western India
Bengaluru	12.97	77.59	920	Southern India
Imphal	24.81	93.93	786	Eastern India
Jaipur	26.91	75.78	431	Northern India
Kolkata	22.57	88.36	9.1	Eastern India
Port Blair	11.62	92.72	16	Southern India
Pune	18.52	73.85	560	Western India
Srinagar	34.08	74.79	1585	Northern India
Trivandrum	8.52	76.93	10	Southern India

Table 1: Geographical information of the nine Indian cities

The data supplied by IMD consisted of the maximum and minimum temperature in a month. We computed the average temperature of a month by averaging the maximum and minimum temperatures for a month [World Meteorological Organisation (2012)]. We treat the 12 monthly average temperature readings for a given year as observations from a random sample path of $GP_{PE}(\mu, v, w, \gamma)$ observed at time points $t = 1, \ldots, 12$. Further, we assume that the 53 random sample paths corresponding to the years 1961-2013 are mutually independent as in Berkes et al. (2009). Since, the dataset supplied by IMD contained some missing values we carried out missing value treatment to obtain the completed dataset which is used for this paper. As an example of the kind of missing value treatment carried out, consider Bengaluru city for which the IMD data has two missing values for March 1962 and August 1963. To impute the missing value for March 1962, a cubic smoothing spline is fitted by taking the available values between December 1961 and January 1963 in R-software. The fitted value for March 1962 is used as the imputed value. Similar procedure is used for imputing the value of August 1963.

The powered exponential covariance function could not be used directly for this data because of the periodic nature of the monthly temperature data. In this context, it is natural to expect that the monthly temperatures of January and December in a year would be correlated because both these months fall in the winter season in India. However this would not be correctly captured by the PECF $\Sigma(s, t) = v_0 \exp(-w|s - t|^{\gamma})$ since the month of January with $s = 1$ would appear to be distant with December with $t = 12$ and therefore the model would expect the temperatures of these months to have low correlation. To overcome the problem we use the following periodic PECF (P^2ECF)$\Sigma(s,t) = v_0 \exp\left(-w\left(2\left|\sin\frac{s-t}{2}\right|\right)^{\gamma}\right)$,which is discussed in a different context in Solin and Särkkä (2014). The motivation behind this formulation is that, if the twelve months in a year are viewed as 12 equispaced points on the unit circle with coordinates $\left(\cos\frac{\pi \text{i}}{6}, \sin\frac{\pi \text{i}}{6}\right), i = 1, \ldots, 12$ then the points representing January and December are adjacent to one another. The Euclidean distance between two points $\left(\cos\frac{\pi \text{s}}{6}, \sin\frac{\pi \text{s}}{6}\right)$ and $\left(\cos\frac{\pi \text{t}}{6}, \sin\frac{\pi \text{t}}{6}\right)$ is $2\left|\sin\frac{s-t}{2}\right|$ which is used to replace $|s-t|$ in the original definition of PECF to get P^2ECF.

We illustrate the data analysis carried out by us using the example of Bengaluru. The data analyses for the other eight cities follow a similar procedure. While the main focus of our investigation is to detect a change in the mean function, if present, we also allow the possibility that the covariance matrix may have changed simultaneously along with the mean i.e. we do not assume that $v_0 = v_1$, $w_0 = w_1$, and $\gamma_0 = \gamma_1$. Since the GLR test for change point problem generally does not perform well if the change point is at the beginning of the given sequence or towards the end, therefore we decided to search for a change point only in the middle of the range. More specifically, we assume that there is no change point in the temperature series for Bengaluru between the years 1961-1970 and also 2004-2013 i.e the range of possible values of r is $11 \leq r \leq 43$. The maximum of value l_0 and l_{1r} , $11 \leq r \leq 43$ are computed by maximizing these using the DEoptim package in R. These are used to compute the value of the GLR test statistic for Bengaluru. The cut-off value of the GLR test statistic is obtained through simulation under the null hypothesis of no change. Since we have assumed that there is no change point in the initial ten years 1961-1970, we compute an estimate of the mean function $\mu(t)$ at the points $t = 1, \ldots, 12$ and the parameters v, w and γ determining the covariance function. The pointwise average of the first ten functional observations is taken as the estimate of estimate of $\mu(t)$. Let s_{uv} be the sample covariance of $X(u)$ and $X(v)$ computed using the first ten observations. We estimate v, w, and γ by minimizing the square of the Frobenius norm, [Golub and Van Loan (1996) (p.55)] of the matrix $(\Sigma(u,v) - s_{uv})_{1 \leq u,v \leq 12}$. We prefer this approach for estimating v, w, and γ because of the small sample size. We treat these estimated values as true values of $\mu_0, v_0, w_0, \gamma_0$ and perform 1000 simulations. For each simulation we generate a random sample of 53 functional observations from this Gaussian process $GP_{P^2E}(\mu_0, v_0, w_0, \gamma_0)$ using the gaussSamp function in the gptk package in R.

The simulated null distribution of the GLR test statistic is obtained by computing the value of the GLR test statistic (with the restriction 11 $\leq$ r $\leq$ 43) for each simulation. In this paper, we reject the null hypothesis of no-change if the observed value of the test statistic is greater than the 95^{th} percentile obtained from the simulated null distribution. While it may be (correctly) argued that these cut-off values are approximate, we have observed that the conclusions drawn on the basis of these cut-off values are mostly in agreement with those using the method of Berkes et al. (2009). Thus we feel that the obtained cut-off values are reasonably robust as far as drawing conclusion about the presence of change point in the data. When the above GLR test procedure is applied to the Bengaluru monthly temperature dataset we get the GLR test statistic value to be 24.93. Since the 95^{th} percentile of the simulated null distribution of the GLR test statistic is 16.56, we reject the null hypothesis of no change and conclude that a change point is present in the dataset. The estimated year of change is 1989 which is the year corresponding to the value of r for which ln $\hat{\lambda}$ is maximum as described in section 4.2. A study of the average temperature in the two periods 1961-1989 and 1990-2013 indicates that the temperature of Bengaluru has risen for all months in the year. The maximum increase has happened for the month of January with average January temperature for the period 1990-2013 being $0.776°C$ higher than that for the period 1961-1989. To quantify the extent of increase and compare the same across the cities we need to compute a distance measure between the two average temperature curves. It may be noted that in mathematics there are several ways of computing the distance between two functions of which the L^p-distances ($p \geq 1$) are most well-known, [Bollabás (1999) (p. 24)]. The L^1-distance between the two average temperature curves observed at time points $t = 1, \ldots, 12$ is $\frac{1}{n}\sum_{t=1}^{12} |\hat{\mu_1}(t) - \hat{\mu_0}(t)|$, the L^2-distance is $\sqrt{\frac{1}{n}\sum_{t=1}^{12} (\hat{\mu_1}(t) - \hat{\mu_0}(t))^2}$ and L^∞-distance is $\max_{1 \leq t \leq 12} |\hat{\mu_1}(t) - \hat{\mu_0}(t)|$ where $\hat{\mu_i}$ is the estimate of μ_i, $i = 0, 1$. When these are computed for Bengaluru, we get the L^1-distance to be 0.470, L^2-distance is 0.478 and the L^∞ distance is 0.776.

When the above GLR test procedure is applied to the monthly temperature dataset of the other eight cities it is found that the null hypothesis of no change is rejected for all the cities except Ahmedabad. Table 2 gives a summary of the results obtained for all the nine cities. Table 4 gives the estimates of the v, w and γ for both the pre and post change periods for all the cities. Since no change point is detected for Ahmedabad only the estimates of v, w and γ for the entire period is given. Table

3 gives the L^1, L^2 and L^∞ distances of the average temperature curves before and after the change point for the eight cities where the GLR test indicated the presence of change point. It is seen that Imphal followed by Srinagar are the two cities where there has been the largest changes in annual temperature patterns as measured by both L^1 and L^2 distances. Further, Srinagar is the city for which there has been the maximum change in temperature for a month as seen by the L^∞ distance. The rise in average temperature for the month of March for Srinagar is the highest rise in monthly temperature amongst all the nine cities included in this study. Figure 1 gives the graphs of the average monthly temperature values before and after the year of change for the eight cities where a change point is detected.

City	95^{th} Percentile	GLR test statistic value	Significant at 95% level	Year of change
Ahmedabad	16.657	7.219	Not Significant	*
Bengaluru	16.56	24.93	Significant	1990
Imphal	17.2	35.649	Significant	1994
Jaipur	5.221	23.9	Significant	1997
Kolkata	5.801	27.034	Significant	1986
Port Blair	5.966	28.282	Significant	2001
Pune	6.087	18.859	Significant	2000
Srinagar	17.212	26.78	Significant	1994
Trivandrum	17.341	31.646	Significant	1994

Table 2: Presence of change point and estimated year of change in temperature in the period 1971-2003 for the nine cities

As mentioned in section 1, Berkes et al. (2009) gives a test for presence of change point for functional observations. We apply this test to the average monthly temperature datasets of all the nine cities. We follow Horváth and Kokoszka (2012) regarding the choice of the number of functional principal components (FPCs) to be used for carrying out this test. As suggested in p.87 of Horváth and Kokoszka (2012), we use the number of FPCs for which 85 % (approx.) of the variability is explained. For all the cities except Imphal and Trivandrum five FPCs are required whereas for these cities four FPCs sufficed. The test rejected the null hypothesis of no change point for all cities except Ahmedabad and Kolkata at 5% level of significance. Thus we see that the results obtained using the methodology of Berkes et al. (2009) agrees with that of the methodology proposed in this paper for all cities except Kolkata. In this context it may be noted that in Berkes et al. (2009) method the covariance function remains the same before and after the change point while we have allowed

City	L^1 distance	L^2 distance	L^∞ distance
Bengaluru	0.470	0.498	0.776
Imphal	1.229	1.256	1.656
Jaipur	0.792	0.937	1.613
Kolkata	0.518	0.577	0.769
Port Blair	0.910	0.947	1.431
Pune	0.345	0.458	0.966
Srinagar	0.913	1.064	1.797
Trivandrum	0.508	0.536	0.710

Table 3: L^1, L^2 and L^∞ distances of the average temperature curves before and after the change point for the eight cities where the GLR test indicated the presence of change point

Cities	Year	v	w	γ
Ahmedabad	1961-2013	0.956	1.295	0.498
Bengaluru	1961-1989	0.314	1.548	0.923
	1990-2013	0.309	1.587	0.575
Imphal	1961-1994	0.988	1.121	0.402
	1995-2013	0.720	1.490	0.487
Jaipur	1961-1997	1.303	1.789	0.496
	1997-2013	1.427	1.938	1.129
Kolkata	1961-1986	0.659	2.393	0.665
	1987-2013	0.513	2.217	0.417
Port Blair	1961-2001	0.517	0.549	0.406
	2002-2013	0.289	1.254	0.735
Pune	1961-2000	0.664	2.350	1.113
	2001-2013	0.559	1.198	0.362
Srinagar	1961-1998	1.678	1.708	0.221
	1999-2013	1.178	3.28	0.339
Trivandrum	1961-1994	0.218	1.133	0.413
	1995-2013	0.236	1.522	0.620

Table 4: Estimates of the parameters v, w and γ for both the pre and post change periods for all the cities

for the possibility that they may be different.

6 Variation of warming with geographical location

In this section, we investigate whether the amount of temperature change in the Indian cities has any relation with geographical location i.e. it's latitude (LAT), longitude (LONG) and altitude (ALT). Since LAT and LONG are in spherical coordinates, therefore first we transform them into Cartesian coordinates by using the following transformation

$$x = R\cos\theta$$
$$y = R\sin\theta\cos\phi$$
$$z = R\sin\theta\sin\phi$$

where θ denotes the latitude and ϕ denotes the longitude, in radians and $R = 6371000$ m is the radius of Earth.

India has four pronounced seasons: summer(S), monsoon(M), autumn(A) and winter(W)(Wikipedia (2007)). The winter season consists of the months of December, January, February and March; Summer season consists of the months of April, May and June; Monsoon season consists of July, August and September; Autumn has the months of October and November. Three indicator variables (s_1, s_2, s_3) are used to represent the four seasons. s_1 is the indicator of monsoon season, s_2 is the indicator of autumn season, and s_3 is the indicator of winter season. Note that the summer season is represented as $s_1 = s_2 = s_3 = 0$. The dataset used in this analysis consists of the x, y, z, ALT, s_1, s_2, s_3 and Δ where Δ is the average temperature of a month in the post-change period minus the same in the pre-change period. A snapshot of the data set consisting of 108 observations is given in table 5. Note that for Ahmedabad, we take $\Delta = 0$ for all the months since no change point is detected for this city. We observe that magnitude of warming (Δ) varies across cities and seasons which motivates us to investigate a relation of Δ with LAT, LONG, ALT and seasons.

We fit a linear regression model with Δ as the response variable and x, y, z, ALT, s_1, s_2, s_3 as predictor variables:

$$\Delta = \alpha + \beta_1 s_1 + \beta_2 s_2 + \beta_3 s_3 + \beta_4\ x + \beta_5\ y + \beta_6\ z + \beta_7 \text{ALT} + \epsilon \qquad (11)$$

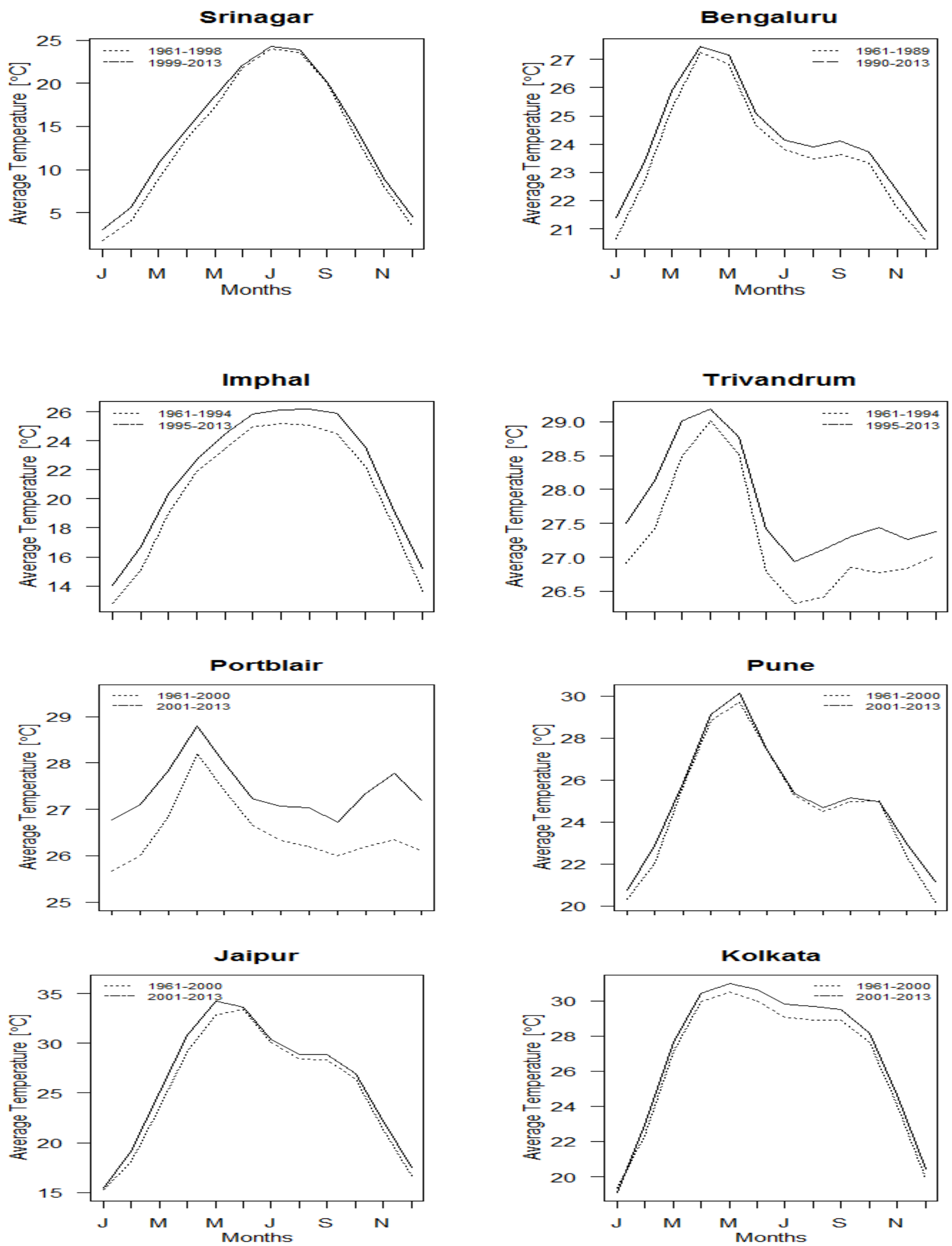

Figure 1: Average temperature curve before change point (dotted line) and the average temperature after change point (smooth solid line) for eight cities where change point is detected

S.no	Δ	s_1	s_2	s_3	x	y	z	ALT
1	0.776	0	0	1	-1429912	-1334234	-6063399	920
2	0.663	0	0	1	-1429912	-1334234	-6063399	920
3	0.632	0	0	1	-1429912	-1334234	-6063399	920
4	0.193	0	0	0	-1429912	-1334234	-6063399	920
5	0.325	0	0	0	-1429912	-1334234	-6063399	920
6	0.439	0	0	0	-1429912	-1334234	-6063399	920
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
104	0.15	1	0	0	-2023657	-1680340	-5802664	560
105	0.16	1	0	0	-2023657	-1680340	-5802664	560
106	-0.03	0	1	0	-2023657	-1680340	-5802664	560
107	0.6	0	1	0	-2023657	-1680340	-5802664	560
108	0.97	0	0	1	-2023657	-1680340	-5802664	560

Table 5: Snapshot of data used to study variation of warming with geographical location, seasons and altitude

where ϵ is the error term satisfying the standard assumptions of a linear regression model. The results of the regression analysis is mentioned in table 6. The value of adjusted R-square for this model is 45.78%.

	Estimate	Std. Error	t value	p-value
(Intercept)	1.05E+01	4.04E+00	2.594	0.01091
s_1	-5.41E-02	9.33E-02	-0.58	0.56334
s_2	1.27E-01	1.04E-01	1.221	0.22479
s_3	2.36E-01	8.73E-02	2.706	0.00802
x	5.16E-07	2.21E-07	2.336	0.02151
y	4.96E-07	8.28E-08	5.988	3.34E-08
z	1.46E-06	6.01E-07	2.433	0.01676
ALT	2.54E-04	8.84E-05	2.866	0.00506

Table 6: Result of regression discussed in section 6

The estimated regression equation indicates that warming is more in those cities which are at high altitude than those in plains. Further, it can be inferred that cities in southern part of India may have witnessed larger amount of warming as compared to cities in northern India having similar altitude. Moreover, the cities in eastern India seems to have warmed more than that of western India with comparable altitudes. Finally, we observe that warming has mostly happened in the Winter season.

In this context it may be noted that popular magazine articles such as Balster (2015) has indicated a higher rise in temperature in regions of higher altitude compared to the plains. Öztürk et al. (2015) discusses the impact of such warming for high altitude ecosystems and suggests that the consequences may be more severe for such regions. Thus the above findings point to an urgent need for putting in place appropriate policy measures that would help in mitigating the effects of warming in the higher altitude regions of India.

7 Conclusion

The present study of monthly temperature data of 53 years from 1961 to 2013 for nine cities in India indicates that temperatures of most cities in India are rising. We come to this conclusion using a new method using Gaussian Process models for analyzing Functional data. The analysis clearly shows that eight of the nine cities located all over India has warmed over the years. Only for Ahmedabad no change point could be detected. The magnitude of warming is seen to be higher for cities located in hilly areas such as Srinagar and Imphal as compared to the cities located in the plains. The findings of this study indicate that there is cause for concern for several Indian cities as the average temperature increase may cause severe impact unless immediate steps are taken to contain or reduce further temperature increase.

References

IPCC (2007). IPCC fourth assessment report. *The Physical Science Basis*, 2:580–595.

Aue, A., Gabrys, R., Horváth, L., and Kokoszka, P. (2009). Estimation of a change-point in the mean function of functional data. *Journal of Multivariate Analysis*, 100(10):2254–2269.

Balster, L. (2015). High-elevation areas may experience the worst of global warming. [Online; accessed 19-Jan-2017].

Berkes, I., Gabrys, R., Horváth, L., and Kokoszka, P. (2009). Detecting changes in the mean of functional observations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(5):927–946.

Bollabás, B. (1999). *Linear Analysis, an introductory course*. Cambridge University Press.

Bosq, D. (2000). *Linear processes in function spaces*. Springer, NewYork.

Erbas, B., Akram, M., Gertig, D. M., English, D., Hopper, J. L., Kavanagh, A. M., and Hyndman, R. (2009). Using functional data analysis models to estimate future time trends in age-specific breast cancer mortality for the United States and England-Wales. *Journal of epidemiology/Japan Epidemiological Association*, 20(2):159–165.

Gao, H. O. and Niemeier, D. A. (2008). Using functional data analysis of diurnal ozone and NOx cycles to inform transportation emissions control. *Transportation Research Part D: Transport and Environment*, 13(4):221–238.

Golub, G. H. and Van Loan, C. F. (1996). *Matrix computations*. The Johns Hopkins University Press.

Horváth, L. and Kokoszka, P. (2012). *Inference for functional data with applications*. Springer.

Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. John Wiley & Sons.

Ikeda, T., Dowd, M., and Martin, J. L. (2008). Application of functional data analysis to investigate seasonal progression with interannual variability in plankton abundance in the Bay of Fundy, Canada. *Estuarine, Coastal and Shelf Science*, 78(2):445–455.

Meiring, W. (2007). Oscillations and time trends in stratospheric ozone levels: a functional data analysis approach. *Journal of the American Statistical Association*, 102(479):788–802.

Mikkonen, S., Laine, M., Mäkelä, H., Gregow, H., Tuomenvirta, H., Lahtinen, M., and Laaksonen, A. (2015). Trends in the average temperature in Finland, 1847–2013. *Stochastic Environmental Research and Risk Assessment*, pages 1–9.

Mullen, K. M. (2014). Continuous Global Optimization in R. *Journal of Statistical Software*, 60(6):1–45.

Müller, H.-G., Wu, S., Diamantidis, A. D., Papadopoulos, N. T., and Carey, J. R. (2009). Reproduction is adapted to survival characteristics across geographically isolated medfly populations. *Proceedings of the Royal Society of London B: Biological Sciences*, pages 4409–4416.

Müller, H.-G. and Yang, W. (2010). Dynamic relations for sparsely sampled Gaussian processes. *Test*, 19(1):1–29.

Öztürk, M., Hakeem, K. R., Faridah-Hanum, I., and Efe, R. (2015). *Climate change impacts on high-altitude ecosystems*. Springer.

Ramsay, J. O. and Dalzell, C. (1991). Some tools for functional data analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 539–572.

Ramsay, J. O., Hooker, G., and Graves, S. (2009). *Functional Data Analysis with R and MATLAB*. Springer.

Ramsay, J. O. and Silverman, B. (2002). *Applied Functional Data Analysis: Methods and case studies*. Springer.

Ramsay, J. O. and Silverman, B. (2005). *Functional data analysis*. Springer.

Shi, J. Q. and Choi, T. (2011). *Gaussian process regression analysis for functional data*. CRC Press.

Silvapulle, M. J. (1996). A test in the presence of nuisance parameters. *Journal of the American Statistical Association*, 91(436):1690–1693.

Smith, D. W. (1978). A simplified approach to the maximum likelihood estimation of the covariance matrix. *The American Statistician*, 32(1):28–29.

Solin, A. and Särkkä, S. (2014). Explicit link between periodic covariance functions and state space models. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*, volume 33, pages 904–912.

Srivastava, M. and Worsley, K. J. (1986). Likelihood ratio tests for a change in the multivariate normal mean. *Journal of the American Statistical Association*, 81(393):199–204.

Storn, R. and Price, K. (1997). Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341–359.

Wahba, G. (1990). *Spline models for observational data*, volume 59. SIAM, Philadelphia.

Wikipedia (2007). Climate of india. [Online; accessed 18-Jan-2017].

World Meteorological Organisation (2012). *WMO Statement on the Status of the Global Climate in 2011*. World Meteorological Association Geneva, Switzerland.

Zamba, K. and Hawkins, D. M. (2006). A multivariate change-point model for statistical process control. *Technometrics*, 48(4):539–549.