Change-Point Analysis of Polar Zone Radiosonde Temperature Data

V. K. JANDHYALA

Department of Mathematics, Washington State University, Pullman, Washington

P. LIU

Philips Healthcare, Seattle, Washington

S. B. FOTOPOULOS

Department of Finance and Management Sciences, Washington State University, Pullman, Washington

I. B. MACNEILL

Department of Statistical and Actuarial Sciences, Western University, London, Ontario, Canada

(Manuscript received 4 March 2013, in final form 18 October 2013)

ABSTRACT

A comprehensive change-point analysis of annual radiosonde temperature measurements collected at the surface, troposphere, tropopause, and lower-stratosphere levels at both the South and North Polar zones has been done. The data from each zone are modeled as a multivariate Gaussian series with a possible change point in both the mean vector as well as the covariance matrix. Prior to carrying out an analysis of the data, a methodology for computing the large sample distribution of the maximum likelihood estimator of the change point is first developed. The Bayesian approach for change-point estimation under conjugate priors is also developed. A simulation study is carried out to compare the maximum likelihood estimator and various Bayesian estimates. Then, a comprehensive change-point analysis under a multivariate framework is carried out on the temperature data for the period 1958–2008. Change detection is based on the likelihood ratio procedure, and change-point estimation is based on the maximum likelihood principle and other Bayesian procedures. The analysis showed strong evidence of change in the correlation between tropopause and lower-stratosphere layers at the South Polar zone subsequent to 1981. The analysis also showed evidence of a cooling effect at the tropopause and lower-stratosphere layers, as well as a warming effect at the surface and troposphere layers at both the South and North Polar zones.

1. Introduction

Identifying climate change through temperature measurements or their proxies continues to play a central role in understanding the dynamics of global climate over time. Changes in temperatures at both surface and upper-atmospheric layers add much clarity toward finding evidence for climate change and global warming. To this end, some recent articles are of interest: McShane and Wyner (2011), Seidel et al. (2011), Thorne and Vose (2010), Thorne et al. (2010), Zou and Wang (2010), Schleip et al. (2009), Randel et al. (2009), Karl et al. (2006), and Ramaswamy et al. (2001). Many of these articles have concluded that changes (cooling or warming) have occurred in the atmospheric temperatures of the earth. These articles take a univariate approach while identifying changes in the mean temperatures as trends on a decadal scale at various atmospheric layers of the globe. In this article, we consider the radiosonde temperature measurements of Angell (2012) at surface, troposphere, tropopause, and lower stratosphere layers of the earth at both the South and North Polar zones for the years 1958–2008. Deviating from the univariate approach, we study annual temperature anomalies at the four layers based on a multivariate model formulation. Mainly, our aim is to take advantage of the multivariate

Corresponding author address: V. K. Jandhyala, Dept. of Mathematics, Washington State University, Pullman, WA 99164-3113.
E-mail: jandhyala@wsu.edu

DOI: 10.1175/JAMC-D-13-084.1

© 2014 American Meteorological Society

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

formulation and identify abrupt changes not only in the means but also in the covariances of temperature anomalies in these four layers.

Sustained changes in climatological factors can occur either gradually or in an abrupt manner. Real changes in climatological factors may be gradual in response to changes in one or more of the external forces known as forcings. However, changes in external forcings that reach beyond a threshold point end up inducing an abrupt change in the climate (Alley et al. 2002; Randall et al. 2007). Abrupt change in climate has been referred to as a large shift in climate that persists (Alley et al. 2002). Such changes may be manifested by changes in average temperature, altered patterns of storms, floods, or droughts over a widespread area. For example, abrupt change can occur as a result of a sudden massive volcanic eruption. Greenhouse gases such as carbon dioxide in the atmosphere can induce abrupt changes in Earth's atmospheric temperatures when they reach beyond a threshold point. Randall et al. (2007) also discuss the possibility of unforced abrupt changes in the climate although such instances might be rare. Apart from such natural abrupt changes, one can also anticipate artificial abrupt changes in climatic data that are a result of changes in data measurement practices (Peterson et al. 1998; Gaffen et al. 2000). In their review article, Peterson et al. (1998) discuss a host of possible ways in which artificial abrupt changes can occur in climatic time series; these include changes in instruments, observing practices, station locations, and formulas used to calculate means. Usually, such artificial abrupt changes can be identified with meticulous documentation of changes in data measurement practices. However, documentation problems in data measurement practices are common and thus artificial abrupt changes in climatic data must also be identified along with real changes. Alley et al. (2002) point out that most statistical models are based on the assumption that climates are not changing but are stationary. Thus, such models have limited use for nonstationary climatic variables that change abruptly over time. They go on to call for statistical approaches that better reflect the properties of abrupt climate changes. This article may be viewed as an attempt to address this issue.

Given the large natural variability in climate time series, it can often be difficult to identify whether and when an abrupt change (real or artificial) may have occurred. From a statistical standpoint, this is often pursued by implementing change detection methods. In simple terms, change-point detection techniques are designed to detect the timing of abrupt changes and decide whether they are significant or not. Thus, there are two inferential problems associated with the identification of the unknown time points at which abrupt changes in the parameters of a statistical model might have occurred: detection and estimation. In the detection part, one aims to develop methods to test for the null hypothesis of no change point in the parameters against the alternative hypothesis of the presence of one or more unknown change points. Upon detecting an unknown change point, one follows it up by estimating its location. While both of these inferential problems are important, the detection part has received greater attention in the literature, including the climatological literature. In this regard, Reeves et al. (2007) give an overview of common change-point detection techniques used in climate literature. Chen and Gupta (2001) provide a survey of different change-point models and Beaulieu et al. (2012) present various applications of change-point detection to climate data. A Bayesian approach has been adapted by Schleip et al. (2009) and Berliner et al. (2000). Ducré-Robitaille et al. (2003) compared eight methods for the detection of shifts in time series data. Overall, change-point methodology has recently become an important tool for identifying changes in climatic factors (Alexandersson 1986; Easterling and Peterson 1995; Jaruskova 1996; Alexandersson and Moberg 1997; Vincent 1998; Horváth et al. 1999; Lund and Reeves 2002; Rodionov 2004; Fearly and Sweeny 2005; Zhao and Chu 2006; DeGaetano 2006; Lau and Wu 2007; Wang et al. 2007; Briggs 2008; Jandhyala et al. 2009; Jaruskova 2010; Fotopoulos et al. 2010).

Upon detecting change in a climatological series, it is important to identify the location of the change point. Knowledge about the location of a change point goes much further than a mere point estimate. One seeks an interval estimate for the change point and this can be constructed via the maximum likelihood estimate (MLE) provided its distribution is known. However, distribution of the change-point MLE is largely intractable under the classical abrupt change modeling (Jandhyala and Fotopoulos 1999; Borovkov 1999; Fotopoulos et al. 2010). Recently, Fotopoulos et al. (2010, 2011) derived computable expressions for the distribution of the change-point MLE when a change occurred in the mean vector of a multivariate Gaussian series and in the parameter of an exponential distribution. In pursuit of performing change-point analysis of radiosonde temperature series observed at four layers of the earth's atmosphere, we first develop in this paper the large sample distribution of the change-point MLE when a change has occurred in both the mean vector and the covariance matrix of a multivariate Gaussian series. Then, we extend the Bayesian conjugate posterior derived by Perreault et al. (2000c) for change in the Gaussian mean vector only to the case of identifying change in both the mean vector as

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

well as the covariance matrix. We also adapt the Bayesian noninformative prior of Son and Kim (2005) and the conditional solution of Cobb (1978) into the change-point analysis of the radiosonde temperature series. This article includes a simulation study for comparing the change-point MLE and three Bayesian solutions; it is also used to evaluate the robustness of the methodologies against deviations from Gaussianity.

The article is organized as follows. A description of the data and the main goal of the study are presented in section 2. Section 3 covers the MLE and different Bayesian approaches for estimating an unknown change point; it also contains a simulation study for evaluating the performance of various methodologies. All aspects of carrying out change-point analysis of the radiosonde series are included in section 4. Section 5 contains a discussion and concluding remarks.

2. Data

Angell (2012) reported radiosonde temperature deviations at various altitudes as a part of the output of 63 globally distributed stations. The data were measured 1) at the surface, 2) in the troposphere (850–300 hPa), 3) at the tropopause (300–100 hPa), and 4) in the lower stratosphere (100–50 hPa). The observations of atmospheric layers were made at both the South Polar (60°–90°S) and North Polar (60°–90°N) zones of the earth. Data on mean annual air temperature measurements for the period 1958–2008 were reported as deviations from the mean of 1958–1977. One may see Angell (2012) for a graphical illustration of the four layers as well as the full details of the sources of the recorded data. Here, we graphically present the data in Figs. 1 and 2 for the South and North Polar zones, respectively.

Our goal in this article is to carry out change-point analyses of the above annual data in order to identify and understand the presence of abrupt changes in the temperature series during the period 1958–2008. While this can be carried out for data from each layer in a univariate manner, we model data from each zone as a multivariate set of dimension $d = 4$. The multivariate formulation enables us to analyze temperature changes at each of the four layers in a simultaneous manner. The ability to analyze temperature changes among the four layers simultaneously is a major advance over the usual univariate methods that detect changes in each layer one at a time. Moreover, the multivariate formulation allows us to identify changes not only in the mean vector but also in the covariance structure, which is otherwise not possible through any of the univariate methods. For detection of change points, we adapt likelihood ratio tests and their distribution theory, as presented by Csörgö and Horváth (1997). However, change-point estimation for changes in a multivariate model requires important developments without which we cannot proceed further. In what follows, we pursue this first for the MLE and then for the Bayesian posterior under the conjugate prior. Other Bayesian procedures are also a part of our study.

3. Change-point estimation in the mean and covariance of a multivariate Gaussian series

We begin by letting $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ be a time series of independent random vectors such that $\mathbf{Y}_i \in R^d$, and for each $i = 1, \ldots, n$, we let $\mathbf{Y}_i$ be multivariate Gaussian with mean vector $\boldsymbol{\mu}$, covariance matrix $\boldsymbol{\Sigma}$, and probability density function (pdf) $f(\cdot\,; \boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then, under the change-point model, we assume $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to have changed from $(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ to $(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ at some unknown change point $\tau_n \in (1, \ldots, n-1)$. In this section, we present results related to the estimation of the unknown change point τ_n under both MLE and Bayesian approaches.

a. Estimation of change point under the MLE approach

Let $\hat{\tau}_n$ be the MLE when $\boldsymbol{\mu}_0$, $\boldsymbol{\mu}_1$, $\boldsymbol{\Sigma}_0$, and $\boldsymbol{\Sigma}_1$ are estimated from data, and $\tilde{\tau}_n$ be the MLE under the known parameters. Since Hinkley (1972) has shown that both $\hat{\tau}_n$ and $\tilde{\tau}_n$ are asymptotically identical, it is sufficient to consider the distribution of $\tilde{\tau}_n$ only. Derivation of the form of the MLE $\tilde{\tau}_n$ under known parameters begins with formulating the likelihood function. Noting that $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ are independent multivariate Gaussian, one expresses the likelihood function for τ_n as

$$L(\tau_n) = \prod_{i=1}^{\tau_n} f(\mathbf{Y}_i; \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \prod_{i=\tau_n+1}^{n} f(\mathbf{Y}_i; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1),$$
$$\tau_n \in (1, \ldots, n-1), \tag{1}$$

where $f(\cdot\,; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ represents the pdf of the multivariate Gaussian distribution. Taking the logarithm on both sides of (1), $\tilde{\tau}_n$ is obtained as the value that maximizes the log-likelihood function. Thus,

$$\tilde{\tau}_n = \text{argmax}_{1 \le j \le n-1} \sum_{i=1}^{j} a(\mathbf{Y}_i), \tag{2}$$

where $a(\mathbf{Y}_i) = \ln[f(\mathbf{Y}_i; \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)/f(\mathbf{Y}_i; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)]$. However, it is convenient to work with the centralized change point estimator, $\xi_n = \tilde{\tau}_n - \tau_n$. Then, it turns out that

$$\xi_n = \text{argmax}_{-\tau_n+1 \le j \le n-\tau_n-1} \sum_{i=1}^{j+\tau_n} a(\mathbf{Y}_i). \tag{3}$$

The above maximizer can be shown to be a result of the following two-sided random walk:

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

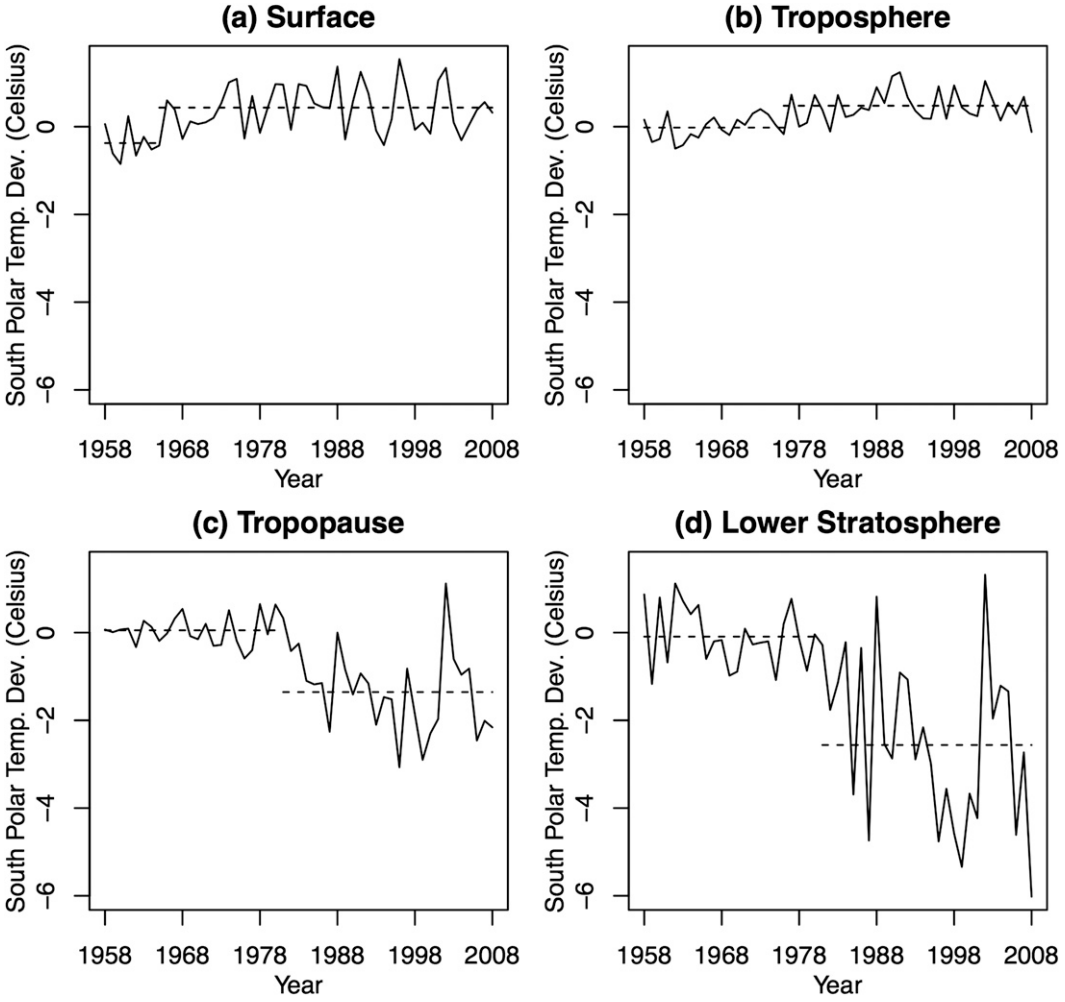

FIG. 1. Annual radiosonde temperature measurement anomalies (°C) for the South Polar zone at the (a) surface, (b) troposphere, (c) tropopause, and (d) lower-stratosphere layers for the years 1958–2008.

$$\Lambda_n(j;\tau_n)=\begin{cases}-\sum_{i=1}^{-j} a(\mathbf{Y}_i^0), & j\in(-\tau_n+1,\ldots,-1)\\ 0, & j=0\\ \sum_{i=1}^{-j} a(\mathbf{Y}_i^*), & j\in(1,\ldots,n-\tau-1),\end{cases} \tag{4}$$

where $\mathbf{Y}_{-i}^0$, $i\in(-\tau_n+1,\ldots,-1)$ are independent and identically distributed (iid) random vectors each with pdf $f(\cdot;\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0)$, and $\mathbf{Y}_i^*$, $i\in(1,\ldots,n-\tau_n-1)$ are iid random vectors each with pdf $f(\cdot;\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)$. Upon letting $X_{-i}^0=-a(\mathbf{Y}_{-i}^0)$, $i\in(-\tau_n+1,\ldots,-1)$, and $X_i^*=-a(\mathbf{Y}_i^*)$, $i\in(1,\ldots,n-\tau_n-1)$, the two random walks in (4) that are independent may be expressed as $S_{-j}^0=\sum_{i=1}^{-j}X_i^0$, $j\in(-\tau_n+1,\ldots,-1)$ and $S_j^*=\sum_{i=1}^{j}X_i^*$, $j\in(1,\ldots,n-\tau_n-1)$.

It is standard practice in change-point literature to let $\tau_n\to\infty$ and $n-\tau_n\to\infty$ while deriving the asymptotic distribution of change-point estimators; one may see Csörgö and Horváth (1997) for a significant amount of literature in this regard. While this is necessary for various technical reasons, its practical significance is that it requires the true change point to stay far enough from either end of the data series. Under this asymptotic setup, it has been shown (see Fotopoulos and Jandhyala 2001) that $\xi_n\to\xi_\infty$ almost surely. This result ensures that the limiting random variable ξ_∞ exists with probability one. One can in principle find the distribution of ξ_∞ by implementing the algorithm derived by Jandhyala and Fotopoulos (1999). The algorithm begins with finding convenient expressions for X^0 and X^*. Details of this and other steps for finding the distribution of ξ_∞ [as expressed in (5) below] are presented in appendix A. In theory, M in (5) is ∞. In reality, one allows M to be suitably large. Often, we found $M=200$ is sufficient:

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

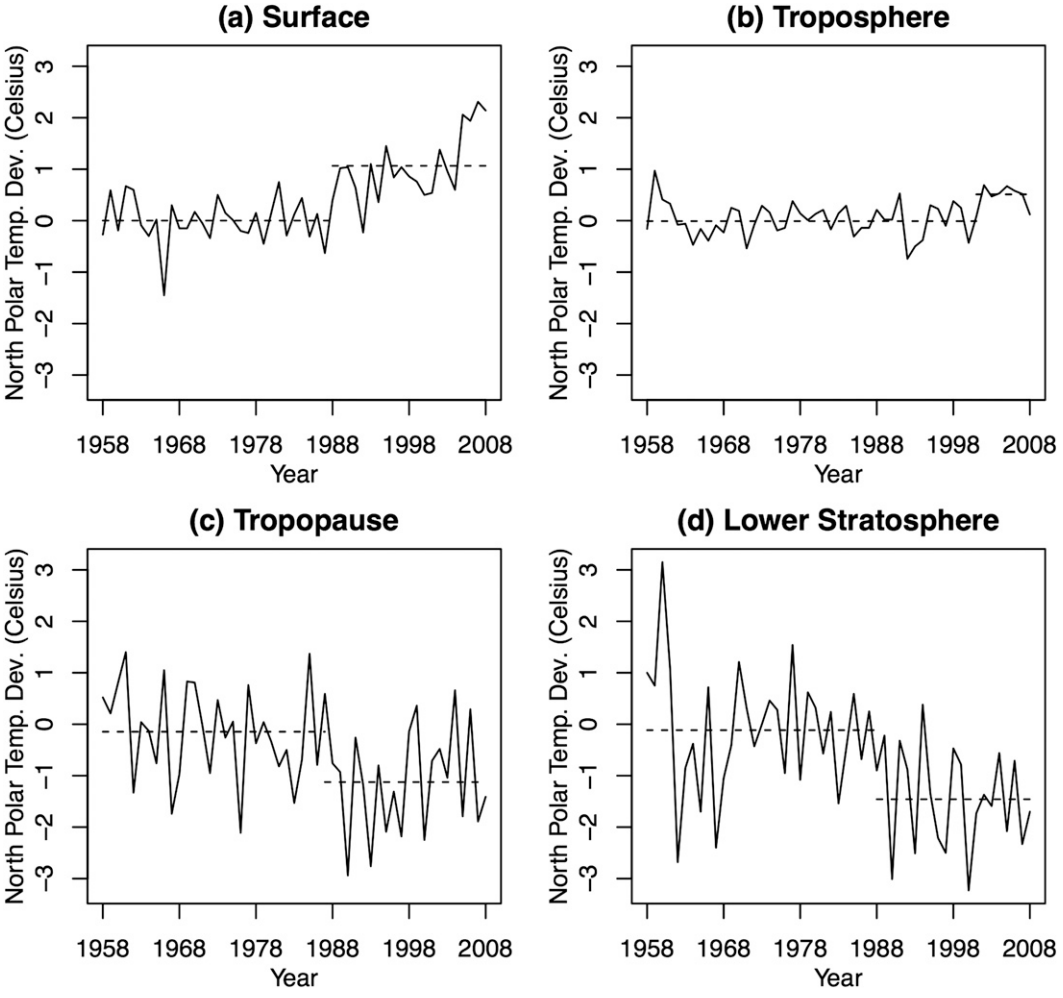

FIG. 2. As in Fig. 1, but for the North Polar zone.

$$P(\xi_\infty = j) \cong \begin{cases} e^{-B^0}[q_j^0 - (1 - e^{-B^*})\tilde{u}_j^0], & j = -M, \quad -M+1, \ldots, -1, \\ e^{-B^0 - B^*}, & j = 0, \\ e^{-B^*}[q_j^* - (1 - e^{-B^0})\tilde{u}_j^*], & j = 1, 2, \ldots, M. \end{cases} \tag{5}$$

b. Estimation of change point under the Bayesian approach

Developments under the Bayesian approach to change-point analysis are vast. Included among them are changes in the mean and/or variance of the univariate normal distribution (Perreault et al. 1999; Perreault et al. 2000a,b), changes in the mean vector of a multivariate normal distribution (Perreault et al. 2000c), changes in linear regression (Seidou et al. 2007), and multiple change points (Fearnhead 2005, 2006; Seidou and Ouarda 2007; Ruggieri 2012).

In this article, we consider three different ways of computing the Bayesian posterior distribution. First, we consider the conjugate priors method that Perreault et al. (2000c) adopted for the case of change in the mean vector only of a multivariate Gaussian sequence and we extend it to the case of change in both the mean vector as well as the covariance matrix. Conjugate priors are convenient in Bayesian analysis due to the fact that both prior and posterior distributions belong to the same family. Next, we state the posterior of the noninformative prior approach adopted by Son and Kim (2005). Finally, we present the conditional solution of Cobb (1978) since

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

it can be applied to the situation of change in both the mean vector and covariance matrix. All three methods, including the derivation of the conjugate priors method, are presented in appendix B.

c. Simulation study

We carry out a simulation study with several goals in mind, all related to the performance of various methods of change-point estimation. The first goal is to assess the closeness of the asymptotic distribution of change-point MLE to the case of finite samples and then to evaluate the asymptotic equivalence of Hinkley (1972) for known and estimated parameters. Our next goal is to carry out a comparison among change-point MLE and the three Bayesian solutions. The final goal is to assess the robustness of the change-point MLE to deviations from normality. Details of the simulation study and its results are all presented in appendix C.

4. Change-point analysis of radiosonde temperatures from South and North Polar zones

a. Detection of change points in radiosonde temperature series

To begin the change detection analyses, let $\mathbf{Y}_1, \ldots, \mathbf{Y}_{51}$ represent four-dimensional data on radiosonde temperatures for a given zone. We then proceed by assuming that $\mathbf{Y}_i$, $i = 1, \ldots, 51$, are serially independent multivariate Gaussian variables with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. We adopt the likelihood ratio method for detecting change points in the temperature series. The statistics and their application are presented in appendix D. In applying the statistics, we take a step-down approach. Beginning with the four-dimensional data, we shall first test the null model of no change [$M^{(0)}$] against the alternative model of change in the mean vector or covariance matrix [$M^{(3)}$]. When this test does not reject the null hypothesis of stationarity, we conclude that no change points exist in the data series. However, when this test is rejected (p value $<$ 0.05), we apply the test to changes in the mean vector alone [$M^{(1)}$] and then to changes in the covariance matrix alone [$M^{(2)}$] upon the residuals from the fit under $M^{(1)}$. We apply this procedure by successively reducing the dimensionality. The p values against all alternatives and for all combinations of layers are presented in Tables 1 and 2 for the South and for North Polar zones, respectively. In these tables, the MLE is reported only when a test is rejected (p value < 0.05), except in one marginal case. Hochberg and Tamhane (1987, theorem 4.1) showed that the step-down procedure of combining outcomes of multiple tests (such as those in Tables 1 or 2) controls the overall significance level.

Applying the step-down procedure to p values in Table 1, we notice that our layer-3,4 bivariate model shows significance against all three alternatives. The corresponding univariate models also show significance, and, moreover, the change-point MLEs coincide (almost) at both univariate and bivariate levels. Thus, at the South Polar zone, we conclude that a change in both the mean and covariance occurred in layers 3 and 4, and we label the corresponding model S-34. Next, we note that the layer-1,2 model shows significance only in the mean. Moving down, the layer-1 and layer-2 models also show significance in their means only. However, the change-point MLE for the layer-1,2 model is 19 whereas the corresponding MLE for the layer-1 model is 8 and it is 19 for the layer-2 model. Because of this discrepancy, we pursue change for the layer-1 (model S-1) and layer-2 (model S-2) models on a univariate basis only. Similarly, applying the same step-down procedure to p values in Table 2 for the North Polar zone, we conclude that a significant change in the mean vector only has occurred in the layer-1,4 two-dimensional model (model N-14). Next, as we step down, layer-2 shows a change in the mean only (p value = 0.0406). The same test for layer-3 shows that the p value = 0.0637, which we conclude to be marginally significant. Thus, we pursue change in the mean only for the layer-2 (model N-2) and layer-3 (model N-3) models on a univariate basis.

b. Cross validation of the assumptions of Gaussianity and independence

The estimation methodologies presented in section 3 cannot proceed without cross validating the assumptions of independence and Gaussianity. We begin our cross-validation process with the data from the South Pole wherein the models to be pursued are S-34, S-1, and S-2. First, let $(\boldsymbol{\mu}_{0,\text{S-34}}, \boldsymbol{\Sigma}_{0,\text{S-34}})$ and $(\boldsymbol{\mu}_{1,\text{S-34}}, \boldsymbol{\Sigma}_{1,\text{S-34}})$ be the respective mean vector and covariance matrix before and after the unknown change point for model S-34. From Table 1, the change-point MLE for model S-34 is 24 or year 1981. Then, estimates of the parameters for the S-34 model enable us to compute the bivariate residuals. These residuals in their standard form can be utilized for model diagnostics. To examine univariate normality, we applied a Shapiro–Wilk test (p values are 0.4155 for layer 3 and 0.1555 for layer 4), and for multivariate normality, we applied skewness and kurtosis tests of Mardia (1970) (p values are 0.8053 and 0.0778) and Henze and Zirkler (1990) (p value 0.1140). Independence is examined from autocorrelation function (ACF) and partial autocorrelation function (PACF) plots each up to lag 10, and there was no significance. Nor were the forward and backward cross correlations significant. Thus, we conclude that there is no evidence against the underlying assumptions

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

TABLE 1. The p values and the corresponding change-point MLEs based on application of the likelihood ratio statistic under alternative hypotheses of change in mean only [$M^{(1)}$], change in covariance only [$M^{(2)}$], and change in both mean and covariance [$M^{(3)}$] for the radiosonde temperature data from the South Polar zone for the years 1958–2008 at atmospheric layers of the surface (1), troposphere (2), tropopause (3), and lower stratosphere (4).

Layers in the model	Mean and covariance/ variance $M^{(3)}$		Mean only $M^{(1)}$		Covariance/ variance only $M^{(2)}$	
	$\hat{\tau}_n$	p value	$\hat{\tau}_n$	p value	$\hat{\tau}_n$	p value
1, 2, 3, 4	25	<0.0001	27	0.0002	26	0.0001
1, 2, 3	25	<0.0001	25	0.0002	4	0.0019
1, 2, 4	27	<0.0001	27	0.0003	28	0.0034
1, 3, 4	24	<0.0001	26	0.0009	24	0.0013
2, 3, 4	25	<0.0001	25	0.0002	24	0.0015
1, 2	14	0.0049	19	0.0156	—	0.3983
1, 3	26	0.0001	26	0.0007	29	0.0108
1, 4	24	0.0001	27	0.0030	28	0.0114
2, 3	25	<0.0001	25	0.0002	18	0.0159
2, 4	27	<0.0001	27	0.0004	27	0.0166
3, 4	24	<0.0001	26	0.0012	24	0.0033
1	8	0.0424	8	0.0291	—	0.5386
2	19	0.0116	19	0.0076	—	0.8357
3	26	0.0001	26	0.0006	29	0.0041
4	27	0.0003	27	0.0019	28	0.0097

of the normality and independence of our model. Similarly, univariate diagnostics for residuals for models S-1 and S-2 showed no evidence against either normality or independence. Finally, similar diagnostics applied to residuals from models N-14, N-2, and N-3 of the North Polar zone showed no evidence against normality or independence.

c. *Parameter estimates and distributions of change-point estimates*

Upon estimating the parameters of any given model for the South or North Polar zones, one computes the asymptotic distribution of the MLE as well as the three Bayesian posteriors of the change point by following the details provided in the respective appendixes. The computed distributions for models S-34, S-1 , and S-2 of the South Polar zone are presented in Tables 3–5. We skip presenting similar tables for models from the North Polar zone to save space. One may then use these distributions to compute a confidence interval estimate of the change point through the distribution of the MLE, and credible regions for the change point through any of the three Bayesian solutions. For each model, the computed parameter estimates, the approximately 95% confidence intervals as well as the approximately 95% Bayesian credible regions, are presented in Tables 6 and 7 for the South and North Polar zones, respectively.

TABLE 2. As in Table 1, but for the North Polar zone.

Layers in model	Mean and covariance/ variance $M^{(3)}$		Mean only $M^{(1)}$		Covariance/ variance only $M^{(2)}$	
	$\hat{\tau}_n$	p value	$\hat{\tau}_n$	p value	$\hat{\tau}_n$	p value
1, 2, 3, 4	44	<0.0001	37	0.0015	5	0.0008
1, 2, 3	31	0.0001	31	0.0016	—	0.1199
1, 2, 4	31	0.0001	31	0.0015	43	0.0634
1, 3, 4	44	<0.0001	37	0.0015	44	0.0131
2, 3, 4	44	0.0001	38	0.0078	10	0.0229
1, 2	31	0.0007	31	0.0013	—	0.2807
1, 3	47	0.0010	31	0.0018	—	0.3055
1, 4	47	0.0007	31	0.0016	—	0.1937
2, 3	32	0.0297	30	0.0397	—	0.9792
2, 4	37	0.0034	37	0.0075	—	0.2457
3, 4	38	0.0056	38	0.0419	12	0.0425
1	31	0.0015	31	0.0011	—	0.1468
2	44	0.0421	44	0.0406	—	0.3762
3	—	0.1215	30	0.0637	—	0.7495
4	32	0.0390	32	0.0222	—	0.1043

d. *Summary of conclusions from change-point models*

The models and the various estimates presented in Tables 6 and 7 enable us to draw important conclusions about the behavior of radiosonde temperature measurements over time at both the South and North Polar zones. First, it is clear from model S-34 and its parameter estimates that the amount of change in absolute (standardized) terms is −1.41°C (−2.01°C) at the tropopause layer and it is −2.47°C (−1.74°C) at the stratosphere layer. In addition, models S-1 and S-2 suggest that the amount of change at the surface layer is 0.81°C (1.65°C) and that at the troposphere layer is 0.50°C (1.55°C). Thus, a cooling effect has occurred around the year 1981 in both the tropopause and stratosphere layers of the South Polar zone whereas there is a warming effect in the surface and troposphere layers with the warming taking place at the surface layer much earlier (around 1965) compared to the year 1976 at the troposphere layer. Moreover, estimates of model S-34 suggest that a significant increase in the variability also took place at the tropopause and stratosphere layers of the South Polar zone. More importantly, the estimated correlation between the tropopause and stratosphere layers before 1981 is −0.07, whereas it has changed to 0.83 after 1981. Thus, a strong positive correlation has begun between these two layers that was nonexistent before 1981.

As for the North Polar zone, estimates for models N-14, N-2, and N-3 suggest that the amount of change at each layer and the year in which the change has occurred are 1.06°C (2.03°C) around the year 1988 for the surface

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

TABLE 3. Distributions of change-point estimates obtained by the methods of MLE, Cobb's conditional solution, Bayesian conjugate prior, and Bayesian noninformative prior for bivariate model S-34 involving layer 3 (tropopause) and layer 4 (lower stratosphere) from data on temperature measurements of the South Polar zone during the years 1958–2008.

			Distribution of change-point estimates			
Obs	Year	Dev	MLE	Cobb	Bayesian (conjugate)	Bayesian (noninformative)
19	1976	−5	0.0000		0.0000	0.0001
20	1977	−4	0.0002	0.0000	0.0001	0.0004
21	1978	−3	0.0008	0.0005	0.0005	0.0011
22	1979	−2	0.0041	0.0030	0.0033	0.0049
23	1980	−1	0.0266	0.0394	0.0261	0.0289
24	1981	0	0.8276	0.7442	0.4859	0.4344
25	1982	1	0.1150	0.0785	0.1055	0.1064
26	1983	2	0.0210	0.1283	0.3267	0.3505
27	1984	3	0.0044	0.0060	0.0519	0.0733
28	1985	4	0.0010	0.0000	0.0000	0.0000
29	1986	5	0.0002		0.0000	0.0000
30	1987	6	0.0001		0.0000	0.0000
Bias			0.1377	0.3061	0.8797	0.9834
RMSE			0.5688	0.7800	1.3877	1.1272

layer, 0.52°C (0.94°C) around the year 2001 at the troposphere layer, −1.27°C (−1.31°C) around the year 2001 for the tropopause layer, and −1.34°C (1.22°C) around the year 1988 for the stratosphere layer. Thus, the changes at both the South and North Polar zones are similar in the sense that the warming effect at the surface and troposphere layers and the cooling effect at the tropopause and stratosphere layers found at the South Polar zone also persist at the North Polar zone. However, they also differ in the sense in that the change in correlation structure found at the tropopause and stratosphere layers of the South Polar zone is not observed at the North Polar zone. Even the variability in temperatures at various layers of the North Polar zone remained steady throughout the time period. Also, the warming and cooling effects began occurring much earlier at the South Polar zone (between the years 1965 and 1981) in comparison to the later year changes (between the years 1987 and 2001) found at the North Polar zone.

e. Change-point models versus linear-trend models

Climatologists often look for linear trends in atmospheric factors. Despite its significance against no change, a change-point model is preferred only when it is significantly better than a linear-trend model. For example, Seidel and Lanzante (2004) compared linear-trend

TABLE 4. As in Table 3, but for univariate model S-1 involving layer 1 (surface).

			Distribution of change-point estimates			
Obs	Year	Dev	MLE	Cobb	Bayesian (conjugate)	Bayesian (noninformative)
3	1960	−5	0.0078	0.0015	0.0039	0.0051
4	1961	−4	0.0139	0.0008	0.0019	0.0025
5	1962	−3	0.0261	0.0076	0.0098	0.0108
6	1963	−2	0.0533	0.0182	0.0181	0.0188
7	1964	−1	0.1288	0.1146	0.0838	0.0794
8	1965	0	0.5272	0.5339	0.3473	0.3054
9	1966	1	0.1288	0.0792	0.0628	0.0606
10	1967	2	0.0533	0.0237	0.0273	0.0286
11	1968	3	0.0261	0.0669	0.0717	0.0712
12	1969	4	0.0139	0.0495	0.0691	0.0707
13	1970	5	0.0078	0.0447	0.0813	0.0846
14	1971	6	0.0046	0.0354	0.0895	0.0953
15	1972	7	0.0027	0.0200	0.0794	0.0881
16	1973	8	0.0017	0.0038	0.0336	0.0412
17	1974	9	0.0010	0.0001	0.0062	0.0091
Bias			0	0.9477	2.4419	2.8884
RMSE			1.8674	2.1457	4.1490	3.9747

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

TABLE 5. As in Table 4, but for univariate model S-2 involving layer 2 (troposphere).

Obs	Year	Dev	Distribution of change-point estimates			
			MLE	Cobb	Bayesian (conjugate)	Bayesian (noninformative)
11	1968	−8	0.0023	0.0031	0.0125	0.0138
12	1969	−7	0.0036	0.0235	0.0561	0.0607
13	1970	−6	0.0058	0.0325	0.0557	0.0579
14	1971	−5	0.0096	0.0803	0.1011	0.1026
15	1972	−4	0.0164	0.0567	0.0589	0.0582
16	1973	−3	0.0293	0.0247	0.0246	0.0241
17	1974	−2	0.0571	0.0192	0.0183	0.0179
18	1975	−1	0.1307	0.0497	0.0387	0.0373
19	1976	0	0.4930	0.3378	0.2154	0.2047
20	1977	1	0.1307	0.0300	0.0258	0.0250
21	1978	2	0.0571	0.0899	0.0692	0.0665
22	1979	3	0.0293	0.1745	0.1381	0.1331
23	1980	4	0.0164	0.0163	0.0210	0.0210
24	1981	5	0.0096	0.0078	0.0139	0.0142
25	1982	6	0.0058	0.0398	0.0631	0.0637
26	1983	7	0.0036	0.0037	0.0118	0.0124
27	1984	8	0.0023	0.0038	0.0155	0.0166
28	1985	9	0.0015	0.0031	0.0176	0.0192
Bias			0	−0.0124	−0.2039	−0.2630
RMSE			2.1144	3.2674	4.6487	4.8115

models against change-point alternatives for modeling temperature series of the atmosphere. Here, we settle this issue by applying the celebrated Akaike information criterion (AIC) procedure that Akaike (1973) proposed for model selection. Omitting details of the computations, we present in Table 8 for the South Polar zone the results of applying the AIC to no change model, to the linear-trend model, and to the proposed change-point model. The same table also contains trend estimates and the corresponding p values that show the nature of their significance. Table 8 shows that the change-point model is the best model by virtue of its AIC value being minimum (as identified by the asterisk) in each case. Similar computations have also been performed for data from the North Polar zone. While we do not present the details of these computations, we have found that the change-point model was the best for all three North Polar zone models also. Overall, it is clear that change-point models explain temperature trends at both the South and North Polar zones better than the models of no change or models with a linear trend.

5. Discussion and concluding remarks

The observed changes at the South and North Polar zones could merely be a result of either instrument changes or changes introduced in the procedures for operating the instruments. In this regard, Gaffen (1994), Gaffen et al. (2000), Lanzante et al. (2003), and Angell (2003) looked into the contamination of data due to changes in radiosonde instruments and the corresponding operational procedures. Angell (2003) pointed out that there were nine stations that had anomalies in measurements and they were all located in the tropical region of the globe. Gaffen et al. (2000) noted changes in instrumentation at stations located in Tahiti, Japan, the former Soviet Union, Australia, and New Zealand. On balance, there is no evidence of documentation or instrumentation problems that may have caused the changes in the data series at both the South and North Polar zones.

The above being the case, what do climatological studies have to say regarding the cooling effect at the tropopause and stratosphere layers and the warming effect at the surface and troposphere layers? Moreover, what are the consequences of these temperature changes?

Climatologists have been investigating changes in atmospheric temperatures for some time (see Angell 1986, 1999; Randel and Wu 1999; Compagnucci et al. 2001; Ramaswamy et al. 2001; Karcher et al. 2003; Comiso 2003; Schleip et al. 2009; Randel et al. 2009; Thorne et al. 2010; Seidel et al. 2011). These studies are based on different datasets that include radiosonde, Microwave Sounding Unit (MSU), rocketsonde, and lidar observations. Yet, their conclusions are similar in that there has been a warming effect at the surface and troposphere layers and a cooling effect at the tropopause and stratosphere layers. Instead of dwelling on the details of these studies, we address below the issue of climatological consequences of the observed temperature changes.

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

TABLE 6. Model parameter estimates and confidence set based on MLE and the three Bayesian credible regions for the model change point from models of the South Polar zone.

Model No.	Point estimate	Model parameter estimates	Confidence set based on MLE and the three Bayesian credible regions
S-34	1981	$\hat{\boldsymbol{\mu}}_{0,\text{S-34}} = (0.0525, -0.0913)^{\text{T}}$ $\hat{\boldsymbol{\mu}}_{1,\text{S-34}} = (-1.3556, -2.5626)^{\text{T}}$	MLE: (1980, 1981, 1982) Cobb: 1979–83
		$\hat{\boldsymbol{\Sigma}}_{0,\text{S-34}} = \begin{pmatrix} 0.1069 & -0.0147 \\ -0.0147 & 0.4329 \end{pmatrix}$	Conjugate: 1978–84
		$\hat{\boldsymbol{\Sigma}}_{1,\text{S-34}} = \begin{pmatrix} 0.8351 & 1.4090 \\ 1.4090 & 3.4279 \end{pmatrix}$	Noninformative: 1978–84
S-1	1965	$\hat{\mu}_{0,\text{S-1}} = -0.3750$ $\hat{\mu}_{1,\text{S-1}} = 0.4347$ $\hat{\sigma}_{\text{S-1}} = 0.4919$	MLE: 1962–68 Cobb: 1964–71 Conjugate: 1964–73 Noninformative: 1963–73
S-2	1976	$\hat{\mu}_{0,\text{S-2}} = -0.0211$ $\hat{\mu}_{1,\text{S-2}} = 0.4769$ $\hat{\sigma}_{\text{S-2}} = 0.3217$	MLE: 1972–79 Cobb: 1969–82 Conjugate: 1969–85 Noninformative: 1968–85

Leaving aside the surface layer, we discuss below existing studies from climatological literature that focus on changes in the troposphere, tropopause, and stratospheric layers of the atmosphere. In this regard recent studies (see Santer et al. 2003; Seidel and Randel 2006; Son et al. 2011; Feng et al. 2012) have found that the tropopause layer plays a central role in how temperature trends vary over time in these three layers. While there may be several other sources in the literature, it seems to us that these four articles are representative of the current knowledge on the coupling effect of the tropopause with the troposphere and stratospheric layers. In this regard, Santer et al. (2003) note that the tropopause represents the boundary between the troposphere and stratosphere and is marked by large changes in the thermal, dynamical, and chemical structures of the atmosphere. Supporting this view, Son et al. (2011) observed that due to its promise as an indicator of climate change, the tropopause has often been examined to better understand stratosphere–troposphere exchange and coupling. Feng et al. (2012) stated that the tropopause has attracted much attention in the recent literature owing to its extreme sensitivity to climate variability and climate change. On the basis of reanalysis data and climate model simulations, recent studies have found that the global tropopause height may be a sensitive indicator of anthropogenic climate change (Seidel and Randel 2006).

The summary conclusion of all these articles has been that the warming of the troposphere and the cooling of the stratosphere contribute significantly to increases in the height of the tropopause. For example, based on results from Parallel Climate Model (PCM) experiments and the reanalysis of data for the period 1979–93 obtained

TABLE 7. As in Table 6, but for the North Polar zone.

Model No.	Point estimate	Parameter estimates	Confidence set based on MLE and the three Bayesian credible regions
N-14	1988	$\hat{\boldsymbol{\mu}}_{0,\text{N-14}} = (0.0013, -0.1152)^{\text{T}}$ $\hat{\boldsymbol{\mu}}_{1,\text{N-14}} = (1.0655, -1.4585)^{\text{T}}$	MLE: 1986–90 Cobb: (1987–89, 1992)
		$\hat{\boldsymbol{\Sigma}}_{\text{N-34}} = \begin{pmatrix} 0.2739 & -0.1535 \\ -0.1535 & 1.2028 \end{pmatrix}$	Conjugate: (1987–89, 1992–94) Noninformative: (1987–89, 1992, 1994)
N-2	2001	$\hat{\mu}_{0,\text{N-2}} = -0.0095$ $\hat{\mu}_{1,\text{N-2}} = 0.5114$ $\hat{\sigma}_{\text{N-2}} = 0.5539$	MLE: 1998–2004 Cobb: 1999–2002 Conjugate: (1994, 1996–2003) Noninformative: 1994–2004
N-3	1987	$\hat{\mu}_{0,\text{N-3}} = -0.1453$ $\hat{\mu}_{1,\text{N-3}} = -1.1243$ $\hat{\sigma}_{\text{N-3}} = 0.9668$	MLE: 1979–95 Cobb: (1979–82, 1985–92) Conjugate: (1961, 1970–75, 1977–92) Noninformative: (1961, 1963, 1964, 1966, 1970–82)

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

TABLE 8. Slope coefficients and the corresponding p values, AIC number for mean only with no change, change-point model with change in mean, and simple linear model with no change for layers 1 and 2; and for bivariate mean vector model with no change, bivariate mean vector with change in mean vector as well as covariance, and bivariate simple linear regression model for layers 3 and 4; for data on temperature measurements from the South Polar zone for the years 1958–2008.

Layer	Model			Residual behavior: Slope (p value)	AIC
1	Mean only with no change			0.0128 (0.0186)	91.98
		Before change		−0.0240 (0.7088)	
1	Change in mean at obs 8				80.36*
		After change		0.0005 (0.9430)	
1	Simple linear regression				88.16
2	Mean only with no change			0.0149 (0.0001)	55.76
		Before change		0.0195 (0.0869)	
2	Change in mean at obs 19				37.03*
		After change		0.0009 (0.8971)	
2	Simple linear regression				39.61
3,4	Bivariate mean vector model with no change		Layer 3	−0.0455 (0.0001)	294.80
			Layer 4	−0.0866	
		Before change	Layer 3	0.0049 (0.4606)	
			Layer 4	−0.0254	
3,4	Bivariate mean vector with change in mean and covariance at obs 24				234.05*
		After change	Layer 3	−0.0321 (0.2867)	
			Layer 4	−0.0779	
3,4	Bivariate linear regression				264.25

from European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA) data, Santer et al. (2003) concluded that both stratospheric cooling and tropospheric warming lead to increases in tropopause height. Seidel and Randel (2006) analyzed radiosonde data from 100 stations archived at the National Climatic Data Center (NCDC) for the period 1980–2004. From their analysis they found that increases in tropopause height may be more directly affected by the cooling of the stratosphere and much less by the warming of the troposphere. In contrast, Feng et al. (2012) carried out an analysis of radiosonde data available in the Integrated Global Radiosonde Archive (IGRA) for the period 1965–2004 and found statistically significant thickening of the tropopause layer. Moreover, they found that the thickness of the tropopause is coupled primarily with the temperature of the lower stratosphere than the upper troposphere.

The conclusions of the present study support the findings of Seidel and Randel (2006) and Feng et al. (2012). It may be recalled that we found the presence of a strong correlation between the tropopause and lower stratosphere (of the South Polar zone) subsequent to 1981 while such a correlation was not present prior to 1981. Also, we have not found any association between the tropopause and troposphere throughout the period of our study (1958–2008). Thus, the warming of the troposphere may not have contributed to the thickness of the tropopause layer whereas the presence of a correlation between the lower stratosphere and tropopause and the cooling of the lower stratosphere appear to have contributed to the increase in the thickness of the tropopause layer. It is noteworthy that the data of Seidel and Randel (2006) begin with 1980, which closely matches with the year in which we found a strong correlation began between the tropopause and lower stratosphere layers. Even Santer et al. (2003) concluded that tropospheric height changes began during 1979.

In the final part of our discussion, we investigate why abrupt changes in the temperatures of atmospheric layers might have occurred. To this end, the anthropogenic and natural forcings studied in the PCM experiments of Santer et al. (2003) have substantial relevance for identifying the causes for tropospheric and stratospheric temperature changes. The anthropogenic forcings that Santer et al. considered were well-mixed greenhouse gases (G), direct scattering effects of surface aerosols (A), and tropospheric and stratospheric ozone (O), while the natural forcings were changes in solar irradiance (S) and volcanic aerosols (V). On the basis of their study, Santer et al. (2003) found that the largest contribution to tropopause height increases during the satellite era (1979–99)

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

came from well-mixed greenhouse gasses and tropospheric and stratospheric ozone. The solar irradiance and volcanic aerosols together reduced the tropopause height, but only marginally.

There are other studies that have regarded ozone depletion since 1980s as the major factor for the cooling of the lower stratosphere (Angell 1986; Randel and Wu 1999; Ramaswamy et al. 2001; Steinbrecht et al. 2003; Cagnazzo et al. 2006). Forster et al. (2007) reasoned that the cooling was due to the decrease in absorption of longwave radiation from reduced ozone levels. Solomon et al. (2007) reported that ozone depletion in the Arctic was far less than that in the Antarctic. Consequently, the cooling at the South Pole was more prominent than in the North Polar zone. Besides ozone depletion, Ramaswamy et al. (2001) found that greenhouse gases such as carbon dioxide not only warmed the surface but also affected the temperatures of the lower stratosphere.

Acknowledgments. We are grateful to the editor and reviewers for constructive comments and suggestions that led to substantial improvements in the paper and its organization. The research of VKJ and SBF was supported by National Science Foundation Grant DMS-0806133. VKJ thanks the C. R. Rao Advanced Institute of Mathematics, Statistics and Computer Science (AIMSCS), Hyderabad, India, for their kind hospitality when he spent time there doing a part of this research work. The research work of IBM was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

APPENDIX A

Asymptotic Distribution of Change-Point MLE

a. Developments toward computational procedure

We adapt the algorithmic procedure of Jandhyala and Fotopoulos (1999) for computing the distribution of the change-point MLE. Thus, the goal is to first find expressions for both X^0 and X^*. Since $\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}_1$ are positive-definite symmetric matrices, they can be decomposed into $\mathbf{\Sigma}_0 = \mathbf{\Sigma}_0^{1/2}\mathbf{\Sigma}_0^{1/2}$ and $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_1^{1/2}\mathbf{\Sigma}_1^{1/2}$, where $\mathbf{\Sigma}_0^{1/2}$ and $\mathbf{\Sigma}_1^{1/2}$ are symmetric. We may then write

$$\mathbf{Y}^0 \overset{D}{=} \boldsymbol{\mu}_0 + \mathbf{\Sigma}_0^{1/2}\mathbf{Z}^0, \quad \text{and} \quad \mathbf{Y}^* \overset{D}{=} \boldsymbol{\mu}_1 + \mathbf{\Sigma}_1^{1/2}\mathbf{Z}^*, \tag{A1}$$

where $\mathbf{Z}^0$ and $\mathbf{Z}^*$ are independent multivariate standard normal variables. Letting $\mathbf{K} = \mathbf{\Sigma}_0^{1/2}\mathbf{\Sigma}_1^{-1/2}$, $\boldsymbol{\eta}^0 = \mathbf{\Sigma}_0^{-1/2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)$, and $\boldsymbol{\eta}^* = \mathbf{\Sigma}_1^{-1/2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)$ and utilizing (A1), one can write X^0 and X^* as

$$\begin{aligned} X^0 = \ln|\mathbf{K}| &- \frac{1}{2}\boldsymbol{\eta}^{0\mathrm{T}}\mathbf{K}\mathbf{K}^{\mathrm{T}}\boldsymbol{\eta}^0 - \frac{1}{2}\boldsymbol{\eta}^{0\mathrm{T}}\mathbf{K}\mathbf{K}^{\mathrm{T}}(\mathbf{I}-\mathbf{K}\mathbf{K}^{\mathrm{T}})^{-1}\mathbf{K}\mathbf{K}^{\mathrm{T}}\boldsymbol{\eta}^0 \\ &+ \frac{1}{2}[\mathbf{Z}^0 + (\mathbf{I}-\mathbf{K}\mathbf{K}^{\mathrm{T}})^{-1}\mathbf{K}\mathbf{K}^{\mathrm{T}}\boldsymbol{\eta}^0]^{\mathrm{T}}(\mathbf{I}-\mathbf{K}\mathbf{K}^{\mathrm{T}}) \\ &\times [\mathbf{Z}^0 + (\mathbf{I}-\mathbf{K}\mathbf{K}^{\mathrm{T}})^{-1}\mathbf{K}\mathbf{K}^{\mathrm{T}}\boldsymbol{\eta}^0] \quad \text{and} \end{aligned} \tag{A2}$$

$$\begin{aligned} X^* = \ln|\mathbf{K}^{-1}| &- \frac{1}{2}\boldsymbol{\eta}^{*\mathrm{T}}\mathbf{K}^{-1}(\mathbf{K}^{-1})^{\mathrm{T}}\boldsymbol{\eta}^* \\ &- \frac{1}{2}\boldsymbol{\eta}^{*\mathrm{T}}\mathbf{K}^{-1}(\mathbf{K}^{-1})^{\mathrm{T}}H(\mathbf{K})\boldsymbol{\eta}^* + \frac{1}{2}[\mathbf{Z}^* - H(\mathbf{K})\boldsymbol{\eta}^*]^{\mathrm{T}} \\ &\times [\mathbf{I}-\mathbf{K}^{-1}(\mathbf{K}^{-1})^{\mathrm{T}}][\mathbf{Z}^0 - H(\mathbf{K})\boldsymbol{\eta}^*], \end{aligned} \tag{A3}$$

where $H(\mathbf{K}) = [\mathbf{I}-\mathbf{K}^{-1}(\mathbf{K}^{-1})^{\mathrm{T}}]^{-1}\mathbf{K}^{-1}(\mathbf{K}^{-1})^{\mathrm{T}}$. Noting that $\mathbf{K}\mathbf{K}^{\mathrm{T}}$ is positive definite, one may write $\mathbf{K}\mathbf{K}^{\mathrm{T}} = \mathbf{\Theta}\mathbf{\Psi}\mathbf{\Theta}^{\mathrm{T}}$, where $\mathbf{\Theta}$ is an orthogonal matrix and $\mathbf{\Psi} = \mathrm{diag}(\psi_1, \ldots, \psi_d)$. Letting $\mathbf{\Theta}^{\mathrm{T}}\boldsymbol{\eta}^0 := \boldsymbol{\omega}^0 = (\omega_1^0, \ldots, \omega_d^0)^{\mathrm{T}}$, and noting that $\mathbf{\Theta}^{\mathrm{T}}\mathbf{Z}^0$ is multivariate standard normal, one can go through various steps to eventually express X^0 as constant plus a weighted sum of independent noncentral chi-square variables each with 1 degree of freedom. Thus,

$$X^0 \overset{D}{=} C^0 + \sum_{s=1}^{d} a_s^0 \chi^2_{1,\sigma_s^{0^2}}, \tag{A4}$$

where $C^0 = (1/2)\ln(\psi_1\psi_2\cdots\psi_d) - (1/2)\sum_{s=1}^{d}[\psi_s\omega_s^{0^2}/(1-\psi_s)]$, $a_s^0 = (1/2)(1-\psi_s)$, the noncentrality parameter $\sigma_s^{0^2} = [\psi_s\omega_s^{0^2}/(1-\psi_s)]^2$, and $s = 1, \ldots, d$. Similarly, X^* in (A3) can also be expressed as

$$X^* \overset{D}{=} C^* + \sum_{s=1}^{d} a_s^* \chi^2_{1,\sigma_s^{*^2}}, \tag{A5}$$

where $C^0 = (1/2)\ln(\psi_1\psi_2\ldots\psi_d) - (1/2)\sum_{s=1}^{d}[\psi_s^{-1}\omega_s^{*2}/(1-\psi_s^{-1})]$, $a_s^* = (1/2)(1-\psi_s^{-1})$, and noncentrality parameter $\sigma_s^{*^2} = [\psi_s^{-1}\omega_s^{*2}/(1-\psi_s^{-1})]$, $s = 1, \ldots, d$. In light of (A4) and (A5), the partial sums S^0_{-j}, $j \in (-\tau_n + 1, \ldots, -1)$, and S_j^*, $j \in (1, \ldots, n - \tau_n - 1)$, may now be expressed as

$$\begin{aligned} &S^0_{-j} = -jC^0 + \sum_{i=1}^{-j}\left(\sum_{s=1}^{d} a^0_{i;s}\chi^2_{i;1,\sigma_s^{0^2}}\right), \\ &j \in (-\tau_n + 1, \ldots, -1) \quad \text{and} \end{aligned} \tag{A6}$$

$$\begin{aligned} &S_j^* = jC^* + \sum_{i=1}^{j}\left(\sum_{s=1}^{d} a^*_{i;s}\chi^2_{i;1,\sigma_s^{*^2}}\right), \\ &j \in (1, \ldots, n - \tau_n - 1), \end{aligned} \tag{A7}$$

where C^0 and C^* are as defined previously, $a^0_{i;s} = a^0_s$, and $\chi^2_{i;1,\sigma_s^{0^2}}$ are independent noncentral chi-square random variables each distributed as $\chi^2_{1,\sigma_s^{0^2}}$, $i \in (1, \ldots, -j)$, $j \in (-\tau_n + 1, \ldots, -1)$, $s = 1, \ldots, d$. Similar observations apply for $a^*_{i;s}$ and $\chi^2_{i;1,\sigma_s^{*^2}}$, $i \in (1, \ldots, j)$, $j \in (1, \ldots, n - \tau_n - 1)$.

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

To compute the distribution of ξ_∞, the above analysis suggests that we require the distribution of weighted sums of independent noncentral chi-square random variables. The number of terms in each of (A6) and (A7) is $|jd|$, $j \in (-\tau_n + 1, \ldots, n - \tau_n - 1)$. Thus, this number increases multiplicatively in terms of n and d. Hence, it is critical that an algorithm is found that can compute probabilities in a fast and efficient manner while maintaining high levels of accuracy and precision (four decimal places) for all values of n and d. In searching the literature, we find that Imhof (1961) developed both exact and approximate methods for computing the distribution of

$$Q = \sum_{s=1}^{d} a_s \chi^2_{h_s,\sigma_s^2}. \tag{A8}$$

Davies (1980) derived an expression of the distribution function for Q. Kuonen (1999) proposed a saddle-point approximation method beginning with the cumulant-generating function $\kappa(\varsigma)$ of Q and assuming that $1 - 2\varsigma a_s > 0$. For saddle-point approximation, one may also see Field (1993). The method of Lu (2006) cannot be used due to a limitation that the sum of degrees of freedom of the chi-square variables could be no greater than 2. In the end, we compared the first three methods through test examples that are taken to be exactly those considered by Imhof (1961). The error in each case was set at 10^{-6}. We skip presenting the results of our computations and only discuss what we have observed. We found that Imhof's exact formula through the integrate function in R, Imhof's approximation, and Davies's method were all equivalent to the true values (up to the fourth decimal). While Kuonen's (1999) saddle-point method was fastest of them all, it was not as accurate and could not be computed for all cases. On balance, we have chosen to implement Imhof's exact formula through the integrate function in R.

b. Implementation of algorithm

We shall now proceed to implement the algorithmic procedure of Jandhyala and Fotopoulos (1999) together with the modifications suggested in Fotopoulos et al. (2010). In implementing the algorithmic procedure of Jandhyala and Fotopoulos (1999), one should keep in mind that we implement here only their second approximation. Consequently, upon completing steps S_0–S_4, we implement steps C_1–C_3 (skipping steps S_5–S_7) by restricting to the second approximation only.

Steps S_0 and S_1: Since steps S_0 and S_1 together require identifying the distributions of X^0 and X^*, clearly this is already accomplished through (A4) and (A5).

Step S_2: Here, one first computes (b_j^0) and (b_j^*), $j = 1, 2, \ldots$, where

$$b_j^0 = P(S_j^0 > 0) = P\left[\sum_{i=1}^{j}\left(\sum_{s=1}^{d} a_{i;s}^0 \chi^2_{i;1,\sigma_s^{0^2}}\right) > -jC^0\right]$$

and

$$b_j^* = P(S_j^* > 0) = P\left[\sum_{i=1}^{j}\left(\sum_{s=1}^{d} a_{i;s}^* \chi^2_{i;1,\sigma_s^{*2}}\right) > -jC^*\right].$$

In practice, one only computes (b_j^0) and (b_j^*) for $j = 1, 2, \ldots, M$, where M is a suitably large finite number. One computes the above probabilities through Imhof's (1961) exact formula and then directly utilizes them to compute $B^0 = \sum_{j=1}^{M} b_j^0/j$ and $B^* = \sum_{j=1}^{M} b_j^*/j$.

Step S_3: Here, one needs to compute ϑ^0 and ϑ^*, both as defined in Jandhyala and Fotopoulos (1999). However, Fotopoulos et al. (2010) have recently shown that both $\vartheta^0 = 1$ and $\vartheta^* = 1$ in all cases, and the same holds here also.

Step S_4: In light of step S_3, here we need to compute $\tilde{b}_j^0 = \tilde{b}_j^0(1)$ and $\tilde{b}_j^* = \tilde{b}_j^*(1)$. Applying integration by parts, it can be shown that $\tilde{b}_j^0 = -F_{S_j^0}(0) + \int_0^\infty e^{-t} F_{S_j^0}(t)\, dt$, and hence $\tilde{b}_j^0$ can be computed numerically through cumulative probabilities for S_j^0, $j = 1, 2, \ldots, M$. Values for $\tilde{b}_j^*$, $j = 1, 2, \ldots, M$ are also computed in a similar way.

Steps C_1–C_3: Upon combining steps C_1–C_3, we proceed to iteratively compute (q_j^0) and (q_j^*), and also $(\tilde{u}_j^0)$ and $(\tilde{u}_j^*)$ in the following manner:

$$q_0^0 = 1, \quad jq_j^0 = \sum_{k=0}^{j-1} b_{j-k}^0 q_k^0; \quad \tilde{u}_0^0 = 1,$$

$$j\tilde{u}_j^0 = \sum_{k=0}^{j-1} \tilde{b}_{j-k}^0 \tilde{u}_k^0, \quad j = 1, 2, \ldots, M \quad \text{and}$$

$$q_0^* = 1, \quad jq_j^* = \sum_{k=0}^{j-1} b_{j-k}^* q_k^*; \quad \tilde{u}_0^* = 1,$$

$$j\tilde{u}_j^* = \sum_{k=0}^{j-1} \tilde{b}_{j-k}^* \tilde{u}_k^*, \quad j = 1, 2, \ldots, M.$$

Last, the asymptotic distribution for the change-point MLE is computed from

$$P(\xi_\infty = j) \cong \begin{cases} e^{-B^0}[q_j^0 - (1 - e^{-B^*})\tilde{u}_j^0], & j = -M, -M+1, \ldots, -1, \\ e^{-B^0 - B^*}, & j = 0, \\ e^{-B^*}[q_j^* - (1 - e^{-B^0})\tilde{u}_j^*], & j = 1, 2, \ldots, M. \end{cases}$$

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

APPENDIX B

Bayesian Approaches to Change-Point Estimation

a. Conjugate prior approach

Let $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ be time series of independent Gaussian vectors, $\mathbf{Y}_i \in R^d$, $i = 1, \ldots, n$, such that $\mathbf{Y}_i \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, $i = 1, \ldots, \tau_n$, $\mathbf{Y}_i \sim N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$, $i = \tau_n + 1, \ldots, n$, and τ_n is unknown. Letting $\mathbf{P}_0 = \boldsymbol{\Sigma}_0^{-1}$ and $\mathbf{P}_1 = \boldsymbol{\Sigma}_1^{-1}$, the following independent prior distributions are assumed: $\boldsymbol{\mu}_0 \sim N[\boldsymbol{\Phi}_0, (\lambda_0 \mathbf{P}_0)^{-1}]$, $\boldsymbol{\mu}_1 \sim N[\boldsymbol{\Phi}_1, (\lambda_1 \mathbf{P}_1)^{-1}]$, $\mathbf{P}_0 \sim$ Wishart$(a_0, \mathbf{B}_0)$, $\mathbf{P}_1 \sim$ Wishart$(a_1, \mathbf{B}_1)$, and $\tau_n \sim$ Discrete Uniform$(1, \ldots, n-1)$. The pdf of $\mathbf{P} \sim$ Wishart$(a, \mathbf{B})$ is given by

$$f(\mathbf{P} \mid a, \mathbf{B}) = \frac{|\mathbf{B}|^{a/2} |\mathbf{P}|^{(a-d-1)/2} \exp[-\frac{1}{2}\mathrm{tr}(\mathbf{BP})]}{2^{ad/2} \Gamma_d\left(\frac{a}{2}\right)}, \tag{B1}$$

where a, the degrees of freedom, satisfies $a > d-1$ and $\Gamma_d(a/2) = \pi^{d(d-1)/4} \prod_{j=1}^{d} \Gamma[(a/2) + (1-j)/2]$. The joint prior distribution of parameters $\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1$ can then be expressed as

$$\begin{aligned} f_{0101}(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1) &= f_{0|0}(\boldsymbol{\mu}_0 \mid \mathbf{P}_0) f_0(\mathbf{P}_0) f_{1|1}(\boldsymbol{\mu}_1 \mid \mathbf{P}_1) f_1(\mathbf{P}_1) \\ &= f_{\mathrm{NWNW}}(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1; \boldsymbol{\Phi}_0, \lambda_0, a_0, \mathbf{B}_0, \boldsymbol{\Phi}_1, \lambda_1, a_1, \mathbf{B}_1) \end{aligned}$$

Since τ_n is independent of $\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1$, the above is also equal to $f_{0101|\tau_n}(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1)$. Then,

$$\begin{aligned} &f(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1 \mid \tau_n, \mathbf{Y}_1, \ldots, \mathbf{Y}_n) \\ &\quad \propto f(\mathbf{Y}_1, \ldots, \mathbf{Y}_n \mid \boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1, \tau_n) f(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1 \mid \tau_n). \end{aligned}$$

Upon carrying out some algebraic simplifications, one obtains

$$\begin{aligned} f(\mathbf{Y}_1, \ldots, \mathbf{Y}_n | \boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1, \tau_n) &= (2\pi)^{-nd/2} |\mathbf{P}_0|^{\tau_n/2} |\mathbf{P}_1|^{(n-\tau_n)/2} \exp\left\langle \frac{-\tau_n}{2} \mathrm{tr}\{\mathbf{P}_0[\mathbf{S}_{1:\tau_n} + (\overline{\mathbf{Y}}_{1:\tau_n} - \boldsymbol{\mu}_0)(\overline{\mathbf{Y}}_{1:\tau_n} - \boldsymbol{\mu}_0)^{\mathrm{T}}]\} \right\rangle \\ &\quad \times \exp\left\langle -\frac{n-\tau_n}{2} \mathrm{tr}\{\mathbf{P}_1[\mathbf{S}_{\tau_n+1:n} + (\overline{\mathbf{Y}}_{\tau_n+1:n} - \boldsymbol{\mu}_1)(\overline{\mathbf{Y}}_{\tau_n+1:n} - \boldsymbol{\mu}_1)^{\mathrm{T}}]\} \right\rangle, \end{aligned} \tag{B2}$$

where $\overline{\mathbf{Y}}_{1:\tau_n} = (1/\tau_n)\sum_{i=1}^{\tau_n} \mathbf{Y}_i$, $\overline{\mathbf{Y}}_{\tau_n+1:n} = [1/(n-\tau_n)]\sum_{i=\tau_n+1}^{n} \mathbf{Y}_i$, $\mathbf{S}_{1:\tau_n} = (1/\tau_n)\sum_{i=1}^{\tau_n} (\mathbf{Y}_i - \overline{\mathbf{Y}}_{1:\tau_n})(\mathbf{Y}_i - \overline{\mathbf{Y}}_{1:\tau_n})^{\mathrm{T}}$, and $\mathbf{S}_{\tau_n+1:n} = \frac{1}{n-\tau_n}\sum_{i=\tau_n+1}^{n} (\mathbf{Y}_i - \overline{\mathbf{Y}}_{\tau_n+1:n})(\mathbf{Y}_i - \overline{\mathbf{Y}}_{\tau_n+1:n})^{\mathrm{T}}$. Consequently, one has

$$\begin{aligned} &f_{\mathrm{NWNW}}(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1; \boldsymbol{\Phi}_0, \lambda_0, a_0, \mathbf{B}_0, \boldsymbol{\Phi}_1, \lambda_1, a_1, \mathbf{B}_1) \\ &\quad = (2\pi)^{-d} |\lambda_0 \lambda_1|^{d/2} |\mathbf{P}_0|^{(a_0-d)/2} |\mathbf{P}_1|^{(a_1-d)/2} \frac{|\mathbf{P}_0|^{a_0/2}}{2^{a_0 d/2} \Gamma_d\left(\frac{a_0}{2}\right)} \frac{|\mathbf{P}_1|^{a_1/2}}{2^{a_1 d/2} \Gamma_d\left(\frac{a_1}{2}\right)} \\ &\quad \times \exp\left\langle -\mathrm{tr}\frac{1}{2}[\lambda_0 \mathbf{P}_0 (\boldsymbol{\mu}_0 - \boldsymbol{\Phi}_0)(\boldsymbol{\mu}_0 - \boldsymbol{\Phi}_0)^{\mathrm{T}} + \mathbf{P}_0 \mathbf{B}_0] - \mathrm{tr}\frac{1}{2}[\lambda_1 \mathbf{P}_1 (\boldsymbol{\mu}_1 - \boldsymbol{\Phi}_1)(\boldsymbol{\mu}_1 - \boldsymbol{\Phi}_1)^{\mathrm{T}} + \mathbf{P}_1 \mathbf{B}_1] \right\rangle. \end{aligned} \tag{B3}$$

One obtains the form of $f(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1 \mid \tau_n, \mathbf{Y}_1, \ldots, \mathbf{Y}_n)$ from (B2) and (B3) as

$$\begin{aligned} f(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1 | \tau_n, \mathbf{Y}_1, \ldots, \mathbf{Y}_n) &\propto |\mathbf{P}_0|^{(a_0'-d)/2} |\mathbf{P}_1|^{(a_1'-d)/2} \\ &\quad \times \exp\left\{ -\mathrm{tr}\frac{1}{2}[\lambda_0' \mathbf{P}_0 (\boldsymbol{\mu}_0 - \boldsymbol{\Phi}_0')(\boldsymbol{\mu}_0 - \boldsymbol{\Phi}_0')^{\mathrm{T}} + \mathbf{P}_0 \mathbf{B}_0'] \right\} \\ &\quad - \mathrm{tr}\frac{1}{2}[\lambda_1' \mathbf{P}_1 (\boldsymbol{\mu}_1 - \boldsymbol{\Phi}_1')(\boldsymbol{\mu}_1 - \boldsymbol{\Phi}_1')^{\mathrm{T}} + \mathbf{P}_1 \mathbf{B}_1'], \end{aligned}$$

where $\lambda_0' = \lambda_0 + \tau_n$, $\lambda_1' = \lambda_1 + n - \tau_n$, $a_0' = a_0 + \tau_n$, $a_1' = a_1 + n - \tau_n$,

$$\boldsymbol{\Phi}_0' = \frac{\lambda_0 \boldsymbol{\Phi}_0 + \tau_n \overline{\mathbf{Y}}_{1:\tau_n}}{\lambda_0 + \tau_n}, \quad \boldsymbol{\Phi}_1' = \frac{\lambda_1 \boldsymbol{\Phi}_1 + (n-\tau_n)\overline{\mathbf{Y}}_{\tau_n+1:n}}{\lambda_0 + n - \tau_n},$$

$$\mathbf{B}_0' = \frac{\tau_n \lambda_0 (\overline{\mathbf{Y}}_{1:\tau_n} - \boldsymbol{\Phi}_0)(\overline{\mathbf{Y}}_{1:\tau_n} - \boldsymbol{\Phi}_0)^{\mathrm{T}}}{\lambda_0 + \tau_n} + \tau_n \mathbf{S}_{1:\tau_n} + \mathbf{B}_0, \quad \text{and}$$

$$\mathbf{B}_1' = \frac{(n-\tau_n)\lambda_1 (\overline{\mathbf{Y}}_{\tau_n+1:n} - \boldsymbol{\Phi}_1)(\overline{\mathbf{Y}}_{\tau_n+1:n} - \boldsymbol{\Phi}_1)^{\mathrm{T}}}{2(\lambda_0 + n - \tau_n)} + \frac{(n-\tau_n)\mathbf{S}_{\tau_n+1:n}}{2} + \mathbf{B}_1.$$

Clearly, the posterior distribution of $(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \mathbf{P}_0, \mathbf{P}_1)$ is NWNW. One then finds $f(\mathbf{Y}_1, \ldots, \mathbf{Y}_n | \tau_n)$ from

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

$$f(\mathbf{Y}_1,\ldots,\mathbf{Y}_n \mid \tau_n) = \frac{f(\mathbf{Y}_1,\ldots,\mathbf{Y}_n \mid \boldsymbol{\mu}_0,\boldsymbol{\mu}_1,\mathbf{P}_0,\mathbf{P}_1,\tau_n) f(\boldsymbol{\mu}_0,\boldsymbol{\mu}_1,\mathbf{P}_0,\mathbf{P}_1 \mid \tau_n)}{f(\boldsymbol{\mu}_0,\boldsymbol{\mu}_1,\mathbf{P}_0,\mathbf{P}_1 \mid \tau_n,\mathbf{Y}_1,\ldots,\mathbf{Y}_n)}.$$

Substituting the respective densities and recognizing that $\tau_n \sim$ Discrete Uniform$(1,\ldots,n-1)$, one applies Bayes's theorem in order to obtain the posterior of τ_n as

$$f(\tau_n \mid \mathbf{Y}_1,\ldots,\mathbf{Y}_n) \propto |\lambda_0'\lambda_1'|^{-d/2} |\mathbf{B}_0'|^{-a_0'} |\mathbf{B}_1'|^{-a_1'} \prod_{k=1}^{d} \Gamma\left(\frac{a_0'+1-k}{2}\right)\Gamma\left(\frac{a_1'+1-k}{2}\right), \tau_n = 1,\ldots,n-1.$$

b. Noninformative prior approach

Son and Kim (2005) adopted Jeffrey's noninformative priors approach with joint prior as

$$\pi(\boldsymbol{\mu}_0,\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_0,\boldsymbol{\Sigma}_1) = c\,(|\boldsymbol{\Sigma}_0||\boldsymbol{\Sigma}_1|)^{-(d+1)/2}, \quad \text{(B4)}$$

where c is the normalizing constant. The prior for τ_n is still set to be discrete uniform on $(1,\ldots,n-1)$. Son and Kim (2005) then derived the joint density of the full sample to be

$$f(\tau_n,\mathbf{Y}_1,\ldots,\mathbf{Y}_n) = c\,\frac{\prod_{i=1}^{d}\Gamma\left(\frac{\tau_n-i}{2}\right)\Gamma\left(\frac{n-\tau_n-i}{2}\right)}{\tau_n^{d/2}(n-\tau_n)^{d/2}\pi^{d(n-d-1)/2}|\mathbf{S}_{1:\tau_n}|^{(\tau_n-1)/2}|\mathbf{S}_{\tau_n+1:n}|^{(n-\tau_n-1)/2}}, \quad \text{(B5)}$$

where $\mathbf{S}_{1:\tau_n}$ and $\mathbf{S}_{\tau_n+1:n}$ are as defined previously. The posterior distribution of the change point can be obtained by a straightforward application of Bayes's theorem.

c. Cobb's conditional approach

Cobb (1978) adopted a conditional MLE approach and suggested that it may be viewed as a Bayesian solution to the change-point problem. Conditioning upon D observations on either side of the MLE $\hat{\tau}_n$, Cobb (1978) provided conditional probabilities for the centered change point as

$$P(\tau_n - \hat{\tau}_n = j \mid \mathbf{Y}_{\hat{\tau}_n - D},\ldots,\mathbf{Y}_{\hat{\tau}_n + D}) \cong \frac{P(\mathbf{Y}_1,\ldots,\mathbf{Y}_n;\hat{\tau}_n = \tau_n + j)}{\sum_{i=-D}^{D} P(\mathbf{Y}_1,\ldots,\mathbf{Y}_n;\hat{\tau}_n = \tau_n + i)}. \quad \text{(B6)}$$

In the above, D should be chosen such that the event $|\tau_n - \hat{\tau}_n| \le D$ occurs with a high probability of $1-\varepsilon(\mathbf{Y},D)$, where the expression for $\varepsilon(\mathbf{Y},D)$ is given in Cobb (1978). One increases D such that $\varepsilon(\mathbf{Y},D)$ becomes less than a threshold, which we set to be 0.0001 in this study.

APPENDIX C

Simulation Study and Results

a. The simulation study

First, we carried out some preliminary simulations and found that the change-point MLE and the three Bayesian solutions for the most part depended on $\boldsymbol{\eta}^0 = \boldsymbol{\Sigma}_0^{-1/2}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0)$, $\boldsymbol{\eta}^* = \boldsymbol{\Sigma}_1^{-1/2}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0)$, and $|\mathbf{K}| = |\boldsymbol{\Sigma}_0^{1/2}\boldsymbol{\Sigma}_1^{-1/2}|$, rather than the actual values of $\boldsymbol{\mu}_0,\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_0,\boldsymbol{\Sigma}_1$. Thus, in designing the simulation study, we only set values for $\boldsymbol{\eta}^0$ and $|\mathbf{K}\mathbf{K}^{\mathrm{T}}|$, and treated $\boldsymbol{\eta}^*$ as being redundant. In the actual simulations, we found it convenient to work with $\delta = \sqrt{\boldsymbol{\eta}^{0\mathrm{T}}\boldsymbol{\eta}^0/2}$. Accordingly, we let without loss of generality $\boldsymbol{\mu}_0 = (0,\ldots,0)^{\mathrm{T}}$ and $\boldsymbol{\Sigma} = \mathbf{I}$; then, selected entries for $\boldsymbol{\mu}_1$ and $\boldsymbol{\Sigma}_1$ to yield a set of preset values for δ and $|\mathbf{K}\mathbf{K}^{\mathrm{T}}|$. In the study, we chose δ to be 1.5 or 2, and let $|\mathbf{K}\mathbf{K}^{\mathrm{T}}|$ be 1, 1.1, or 1.6, with 1 corresponding to the case of change in the mean only. Also, we chose sample size and change-point location combinations to be n/τ: 100/50, 100/30, 50/25. Robustness was assessed by letting the observations follow the multivariate t distribution with degrees of freedom selected to be 5, 10, 20, or infinite. Finally, parameters of the prior distributions under the Bayesian conjugate priors method were chosen appropriately. For any given set of parameters, we performed 100 000 simulations for data generation.

The closeness of any pair of distributions was determined by comparing the means and root-mean-square errors (RMSEs) for the two distributions. The empirical distribution of change-point MLE and Cobb's solution were carried out when $\boldsymbol{\mu}_0,\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_0,\boldsymbol{\Sigma}_1$ were known (k) or estimated (e), through four scenarios: (i) kk, means and covariances known; (ii) ke, means known and covariances estimated; (iii) ek, means estimated and covariances known; and (iv) ee, means and covariances unknown. We adapted the following procedure for implementing these four cases. Let the jth sample generated be $\mathbf{Y}_{j,1},\ldots,\mathbf{Y}_{j,n}$,

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

TABLE C1. RMSE for various change-point estimates in the bivariate case when $n/\tau = 100/50$.

δ	$\lvert\mathbf{K}\mathbf{K}^{\mathrm{T}}\rvert$	Degrees of freedom	MLE theory	MLE kk	MLE ke	MLE ek	MLE ee	Cobb kk	Cobb ke	Cobb ek	Cobb ee	Bayesian Noninformative	Bayesian Conjugate
1.5	1	5		0.9388	0.9366	1.0240	0.9894	0.9988	1.1293	1.0511	1.1965	1.2640	1.2483
1.5	1	10		0.6711	0.6779	0.7030	0.7080	0.7555	0.8191	0.7928	0.8572	0.8857	0.8731
1.5	1	20		0.5780	0.5821	0.6021	0.6069	0.6749	0.7119	0.7063	0.7416	0.7632	0.7527
1.5	1	∞	0.5057	0.5061	0.5109	0.5255	0.5333	0.6146	0.6294	0.6420	0.6542	0.6710	0.6620
1.5	1.1	5		1.0912	1.2630	1.2672	1.7049	1.1524	1.4573	1.2628	1.6658	2.2159	1.3959
1.5	1.1	10		0.7617	0.8156	0.8045	0.8823	0.8616	0.9870	0.9120	1.0575	1.1134	0.9999
1.5	1.1	20		0.6627	0.6973	0.6968	0.7463	0.7759	0.8533	0.8187	0.9062	0.9421	0.8693
1.5	1.1	∞	0.5842	0.5815	0.6084	0.6080	0.6396	0.7086	0.7519	0.7444	0.7872	0.8167	0.7694
1.5	1.6	5		1.2085	1.3068	1.4383	1.7290	1.2595	1.5213	1.4306	1.7334	2.2245	1.4438
1.5	1.6	10		0.8012	0.8545	0.8499	0.9256	0.9089	1.0438	0.9636	1.1209	1.1743	1.0479
1.5	1.6	20		0.7033	0.7424	0.7419	0.7885	0.8249	0.9107	0.8717	0.9657	1.0027	0.9212
1.5	1.6	Inf	0.6279	0.6277	0.6554	0.6567	0.6895	0.7636	0.8096	0.8058	0.8514	0.8797	0.8216
2	1	5		0.5288	0.5278	0.5618	0.5442	0.5510	0.6173	0.5678	0.6365	0.6538	0.6453
2	1	10		0.3509	0.3544	0.3624	0.3656	0.3902	0.4238	0.4034	0.4358	0.4445	0.4389
2	1	20		0.2924	0.2968	0.2993	0.3041	0.3385	0.3601	0.3487	0.3693	0.3750	0.3702
2	1	∞	0.2398	0.2397	0.2444	0.2481	0.2517	0.2952	0.3052	0.3051	0.3130	0.3178	0.3136
2	1.1	5		0.6156	0.6886	0.6645	0.7994	0.6388	0.7941	0.6688	0.8511	0.9841	0.7956
2	1.1	10		0.4067	0.4337	0.4196	0.4506	0.4546	0.5225	0.4701	0.5410	0.5580	0.5322
2	1.1	20		0.3418	0.3641	0.3502	0.3753	0.3980	0.4442	0.4111	0.4577	0.4684	0.4517
2	1.1	∞	0.2873	0.2895	0.3052	0.2975	0.3162	0.3538	0.3813	0.3660	0.3936	0.4013	0.3889
2	1.6	5		0.6854	0.7272	0.7661	0.8289	0.7010	0.8514	0.7587	0.9180	1.0508	0.8484
2	1.6	10		0.4423	0.4693	0.4558	0.4863	0.4964	0.5699	0.5151	0.5927	0.6102	0.5803
2	1.6	20		0.3765	0.3972	0.3857	0.4089	0.4392	0.4887	0.4543	0.5047	0.5171	0.4968
2	1.6	∞	0.3216	0.3238	0.3401	0.3323	0.3526	0.3964	0.4258	0.4107	0.4396	0.4481	0.4332

$j = 1, \ldots, 100\,000$. The four cases of kk, ke, ek, and ee have been implemented by estimating the parameters before the change point as

$$\boldsymbol{\mu}_{0,j}^{\mathrm{ek};t} = \boldsymbol{\mu}_{0,j}^{\mathrm{ee};t} = \frac{1}{t}\sum_{i=1}^{t}\mathbf{Y}_{j,i};$$

$$\boldsymbol{\Sigma}_{0,j}^{\mathrm{ke};t} = \frac{1}{t}\sum_{i=1}^{t}(\mathbf{Y}_{j,i} - \boldsymbol{\mu}_0)(\mathbf{Y}_{j,i} - \boldsymbol{\mu}_0)^{\mathrm{T}};$$

$$\boldsymbol{\Sigma}_{0;t}^{j;\mathrm{ee}} = \frac{1}{t}\sum_{i=1}^{t}(\mathbf{Y}_{j,i} - \boldsymbol{\mu}_{0,j}^{\mathrm{ee};t})(\mathbf{Y}_{j,i} - \boldsymbol{\mu}_{0,j}^{\mathrm{ee};t})^{\mathrm{T}},$$

$t = d+1, \ldots, n-(d-1)$, where d is the dimensionality. Similar expressions have been used for estimating parameters after the unknown change point. Then, the log-likelihood is given by

$$l_j^{\mathrm{est}}(t) = \sum_{i=1}^{t}\ln f(\mathbf{Y}_{j,i}; \boldsymbol{\mu}_{0,j}^{\mathrm{est};t}, \boldsymbol{\Sigma}_{0,j}^{\mathrm{est};t}) + \sum_{i=t+1}^{n}\ln f(\mathbf{Y}_{j,i}; \boldsymbol{\mu}_{1,j}^{\mathrm{est};t}, \boldsymbol{\Sigma}_{1,j}^{\mathrm{est};t})$$

for $j = 1, \ldots, 100\,000$, and $t = d+1, \ldots, n-(d-1)$, and where est assumes one of the (kk, ke, ek, ee). The change-point MLE is obtained from $\hat{\tau}_j^{\mathrm{est}} = \mathrm{argmax}_{d+1\le t\le n-d}[l_j^{\mathrm{est}}(t)]$, est $\in$ (kk, ke, ek, ee).

Values for the bias and RMSE for the change-point MLE have been computed from

$$\mathrm{Bias}(\hat{\tau}^{\mathrm{MLE;est}}) = \frac{1}{N}\sum_{j=1}^{N}(\hat{\tau}_j^{\mathrm{est}} - \tau) \quad \text{and}$$

$$\mathrm{RMSE}(\hat{\tau}^{\mathrm{MLE;est}}) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\hat{\tau}_j^{\mathrm{est}} - \tau)^2},$$

where $N = 100\,000$ and est $\in$ (kk, ke, ek, ee). After computing posterior probabilities for each simulation under the three Bayesian methods, the respective bias and RMSE values have been computed as

$\mathrm{Bias}(\hat{\tau}) = 1/N\sum_{j=1}^{N}\sum_{t=1}^{n-1}(t-\tau)p_{j,t}$ and $\mathrm{RMSE}(\hat{\tau}) = \sqrt{1/N\sum_{j=1}^{N}\sum_{t=1}^{n-1}(t-\tau)p_{j,t}}$, where $p_{j,t}$ is the probability that the change point is at t for the jth simulated sample, $j = 1, \ldots, N$, $N = 100\,000$.

b. Results of the simulation study

First, we present here results for the bivariate case only; the univariate case showed similar results. Next, results for the bias were negligible in all cases and hence we do not present them here. However, RMSE values were found to vary considerably together with the method of estimation, sample size, location of the change point, and amount of change, as well as departure from Gaussianity. Thus, we present them in Tables C1–C3. To better capture the

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

TABLE C2. As in Table C1, but for $n/\tau = 100/30$.

δ	$\lvert\mathbf{K}\mathbf{K}^{\mathrm{T}}\rvert$	Degrees of freedom	MLE theory	MLE				Cobb				Bayesian	
				kk	ke	ek	ee	kk	ke	ek	ee	Noninformative	Conjugate
1.5	1	5		0.9492	0.9442	1.0796	1.0541	1.0065	1.1312	1.0843	1.2171	1.3207	1.3086
1.5	1	10		0.6667	0.6732	0.7034	0.7133	0.7504	0.8163	0.7934	0.8609	0.8956	0.8812
1.5	1	20		0.5782	0.5847	0.6067	0.6168	0.6739	0.7125	0.7089	0.7473	0.7737	0.7618
1.5	1	∞	0.5057	0.5057	0.5116	0.5313	0.5377	0.6138	0.6292	0.6448	0.6568	0.6770	0.6670
1.5	1.1	5		1.1048	1.4299	1.3423	2.7625	1.1544	1.5983	1.2983	1.8653	3.4017	4.0741
1.5	1.1	10		0.7616	0.8438	0.8140	0.9526	0.8594	1.0091	0.9165	1.0955	1.1854	2.1246
1.5	1.1	20		0.6628	0.7181	0.6995	0.7960	0.7744	0.8684	0.8197	0.9290	0.9807	1.6063
1.5	1.1	∞	0.5842	0.5839	0.6191	0.6162	0.6623	0.7097	0.7606	0.7485	0.8045	0.8455	1.2701
1.5	1.6	5		1.2205	1.4770	1.7063	2.7452	1.2543	1.6705	1.6252	1.9364	3.4062	3.9952
1.5	1.6	10		0.8066	0.8875	0.8571	1.0123	0.9112	1.0676	0.9703	1.1589	1.2509	2.1124
1.5	1.6	20		0.7067	0.7573	0.7458	0.8322	0.8270	0.9222	0.8755	0.9848	1.0374	1.6171
1.5	1.6	∞	0.6279	0.6280	0.6652	0.6596	0.7114	0.7639	0.8158	0.8057	0.8644	0.9034	1.2958
2	1	5		0.5325	0.5312	0.5719	0.5498	0.5579	0.6201	0.5773	0.6429	0.6662	0.6570
2	1	10		0.3535	0.3571	0.3636	0.3682	0.3920	0.4259	0.4050	0.4388	0.4480	0.4420
2	1	20		0.2923	0.2969	0.3028	0.3061	0.3387	0.3605	0.3511	0.3709	0.3781	0.3730
2	1	∞	0.2398	0.2397	0.2442	0.2489	0.2540	0.2953	0.3052	0.3068	0.3148	0.3199	0.3154
2	1.1	5		0.6289	0.7853	0.7124	1.2784	0.6384	0.8822	0.6745	0.9857	1.5121	1.7881
2	1.1	10		0.4099	0.4405	0.4238	0.4671	0.4562	0.5279	0.4729	0.5546	0.5786	0.7616
2	1.1	20		0.3436	0.3673	0.3532	0.3828	0.3989	0.4482	0.4130	0.4649	0.4812	0.5821
2	1.1	∞	0.2873	0.2893	0.3066	0.2990	0.3203	0.3544	0.3843	0.3680	0.3980	0.4093	0.4711
2	1.6	5		0.6854	0.7272	0.7661	0.8289	0.7010	0.8514	0.7587	0.9180	1.0508	1.8565
2	1.6	10		0.4423	0.4693	0.4558	0.4863	0.4964	0.5699	0.5151	0.5927	0.6102	0.8121
2	1.6	20		0.3765	0.3972	0.3857	0.4089	0.4392	0.4887	0.4543	0.5047	0.5171	0.6253
2	1.6	∞	0.3216	0.3244	0.3409	0.3352	0.3577	0.3969	0.4268	0.4121	0.4422	0.4546	0.5132

behavior of RMSE values, we also presented them in figures for the following special cases: (a) comparing the behavior of MLE and Cobb's methods, (b) effects of sample size and change-point location, (c) comparison among various methods, and (d) effects of degrees of freedom. However, due to limitation of space we omit all figures, and only present a summary of what we observed in them.

(a) *Comparing MLE and Cobb's methods*: RMSE values under the MLE method were all close to each other, except when $n/\tau = 50/25$, estimation was under ee, $\delta = 1.5$, and $|\mathbf{K}\mathbf{K}^{\mathrm{T}}| = 1.3, 1.6$. Also, there was greater variability under Cobb's method when $\delta = 1.5$ compared with $\delta = 2.0$.

(b) *Effect of sample size and change-point location*: The MLE and Cobb's methods tended to yield higher values under ee and when the sample size was small. When $\delta = 1.5$, and the sample size small, the noninformative prior showed higher RMSE values when the change point was to the middle, while values under conjugate prior were higher when the change point was away from the center.

(c) *Comparison among various methods*: All estimation methods showed similar results when the change in mean was at a higher level ($\delta = 2.0$). However, when the amount of change was small, methods of MLE and Cobb yielded smaller RMSE values than did either of the two Bayesian methods.

(d) *Effect of degrees of freedom*: RMSE values for the MLE were close to each other except at 5 degrees of freedom; other methods also exhibited similar behavior. When the degrees of freedom was 5, the noninformative prior method consistently showed higher RMSE values, particularly so when $\delta = 1.5$. Relatively, the MLE and Cobb's methods produced closer and smaller RMSE values. However, when estimation was ee and the change in the mean was small, conjugate prior showed smaller RMSE values at five degrees of freedom.

Overall, it may be concluded that the MLE method was more robust to changing conditions than were the other methods. Generally, RMSE values for Cobb's method were closer to the MLE method. The two Bayesian methods showed similar values only when $\delta = 2.0$.

APPENDIX D

Likelihood Ratio Change-Detection Statistics

While applying the likelihood ratio method, one may test the null hypothesis of no change in $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ $[M^{(0)}]$ against three alternatives, namely, the change in mean $\boldsymbol{\mu}$ only $[M^{(1)}]$, change in covariance $\boldsymbol{\Sigma}$ only $[M^{(2)}]$, and change in both mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ $[M^{(3)}]$. The derivations of all three test statistics and their asymptotic

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

TABLE C3. As in Table C1, but for $n/\tau = 50/25$.

δ	$\lvert\mathbf{K}\mathbf{K}^{\mathrm{T}}\rvert$	Degrees of freedom	MLE theory	MLE kk	MLE ke	MLE ek	MLE ee	Cobb kk	Cobb ke	Cobb ek	Cobb ee	Bayesian Noninformative	Bayesian Conjugate
1.5	1	5		0.9382	0.9338	1.1072	0.1064	0.9979	1.1340	1.1307	1.2832	1.4983	1.4058
1.5	1	10		0.6622	0.6770	0.7264	0.7499	0.7475	0.8323	0.8238	0.9166	0.9883	0.9446
1.5	1	20		0.5741	0.5873	0.6246	0.6439	0.6710	0.7274	0.7364	0.7948	0.8447	0.8113
1.5	1	∞	0.5057	0.5014	0.5153	0.5436	0.5620	0.6108	0.6436	0.6686	0.6982	0.7351	0.7106
1.5	1.1	5		1.0760	1.3205	1.3277	2.7627	1.1420	1.5186	1.3353	1.8096	3.4087	1.3708
1.5	1.1	10		0.7516	0.8748	0.8408	1.6060	0.8526	1.0582	0.9572	1.2248	1.8857	1.0305
1.5	1.1	20		0.6594	0.7409	0.7233	1.2694	0.7718	0.9117	0.8586	1.0449	1.4520	0.9108
1.5	1.1	∞	0.5842	0.5805	0.6472	0.6336	0.9930	0.7064	0.8027	0.7831	0.9196	1.1392	0.8115
1.5	1.6	5		1.1577	1.3607	1.4447	2.7119	1.2203	1.5775	1.4471	1.8698	3.3626	1.4069
1.5	1.6	10		0.7977	0.9168	0.8879	1.6210	0.9036	1.1176	1.0136	1.2913	1.9197	1.0740
1.5	1.6	20		0.6971	0.7855	0.7693	1.2914	0.8205	0.9708	0.9177	1.1125	1.5082	0.9575
1.5	1.6	∞	0.6279	0.6213	0.6869	0.6798	1.0266	0.7585	0.8574	0.8466	0.9894	1.2032	0.8604
2	1	5		0.5252	0.5226	0.5610	0.5530	0.5510	0.6227	0.5813	0.6644	0.7188	0.6882
2	1	10		0.3470	0.3562	0.3658	0.3776	0.3873	0.4346	0.4104	0.4586	0.4775	0.4618
2	1	20		0.2897	0.2982	0.3064	0.3168	0.3355	0.3689	0.3566	0.3888	0.4021	0.3891
2	1	∞	0.2398	0.2390	0.2458	0.2507	0.2603	0.2932	0.3138	0.3116	0.3303	0.3408	0.3301
2	1.1	5		0.6056	0.7177	0.6837	1.1000	0.6366	0.8324	0.6982	0.9473	1.3275	0.8185
2	1.1	10		0.4038	0.4627	0.4280	0.5754	0.4514	0.5628	0.4824	0.6120	0.6939	0.5702
2	1.1	20		0.3405	0.3803	0.3602	0.4725	0.3954	0.4740	0.4224	0.5120	0.5610	0.4849
2	1.1	∞	0.2873	0.2867	0.3179	0.3042	0.3720	0.3517	0.4080	0.3767	0.4385	0.4648	0.4183
2	1.6	5		0.6650	0.7581	0.7799	1.1525	0.6971	0.8927	0.7871	1.0314	1.4060	0.8701
2	1.6	10		0.4390	0.4992	0.4666	0.6176	0.4935	0.6124	0.5288	0.6715	0.7585	0.6049
2	1.6	20		0.3758	0.4171	0.3964	0.5094	0.4374	0.5225	0.4691	0.5672	0.6187	0.5210
2	1.6	∞	0.3216	0.3211	0.3526	0.3414	0.4037	0.3945	0.4535	0.4241	0.4899	0.5199	0.4637

distributions may be found in Csörgö and Horváth (1997). The twice log-likelihood ratio test statistic for testing the null change model $M^{(0)}$ against the change point model $M^{(i)}$ is given by

$$U_n^{(i)} = \max_{d+1\le t\le n-d-1} U_{n,t}^{(i)}, \quad i = 1,2,3, \qquad \text{(D1)}$$

where

$$U_{n,t}^{(1)} = -n\ln|\hat{\boldsymbol{\Sigma}}_t^{(1)}| + n\ln|\hat{\boldsymbol{\Sigma}}_{0,n}|;$$

$$U_{n,t}^{(i)} = -t\ln|\hat{\boldsymbol{\Sigma}}_{0,t}^{(i)}| - (n-t)\ln|\hat{\boldsymbol{\Sigma}}_{1,t}^{(i)}| + n\ln|\hat{\boldsymbol{\Sigma}}_{0,n}|, i = 2,3;$$

$$\hat{\boldsymbol{\Sigma}}_{0,n} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{0,n})(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{0,n})^{\mathrm{T}};$$

$$\hat{\boldsymbol{\Sigma}}_t^{(1)} = \frac{1}{n}\left[\sum_{i=1}^{t}(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{0,t})(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{0,t})^{\mathrm{T}} + \sum_{i=t+1}^{n}(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{1,t})(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{1,t})^{\mathrm{T}}\right];$$

$$\hat{\boldsymbol{\Sigma}}_{0,t}^{(2)} = \frac{1}{n}\sum_{i=1}^{t}(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{0,n})(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{0,n})^{\mathrm{T}};$$

$$\hat{\boldsymbol{\Sigma}}_{1,t}^{(2)} = \frac{1}{n-t}\sum_{i=t+1}^{n}(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{0,n})(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{0,n})^{\mathrm{T}};$$

$$\hat{\boldsymbol{\Sigma}}_{0,t}^{(3)} = \frac{1}{t}\sum_{i=1}^{t}(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{0,t})(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{0,t})^{\mathrm{T}};$$

$$\hat{\boldsymbol{\Sigma}}_{1,t}^{(3)} = \frac{1}{n-t}\sum_{i=t+1}^{n}(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{1,t})(\mathbf{Y}_i - \hat{\boldsymbol{\mu}}_{1,t})^{\mathrm{T}};$$

$$\hat{\boldsymbol{\mu}}_{0,n} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{Y}_i; \quad \hat{\boldsymbol{\mu}}_{0,t} = \frac{1}{t}\sum_{i=1}^{t}\mathbf{Y}_i;$$

$$\hat{\boldsymbol{\mu}}_{1,t} = \frac{1}{n-t}\sum_{i=t+1}^{n}\mathbf{Y}_i; \quad t = d+1,\ldots,n-d-1.$$

The asymptotic distribution of the statistic in (D1) is obtained by considering an appropriate function of the statistic and then centralizing it. Basically, one considers the following:

$$W_n^{(i)} = [2\ln\ln n\, U_n^{(i)}]^{1/2} - \left\langle 2\ln\ln n + \frac{p^{(i)}}{2}\ln\ln\ln n - \ln\left\{\Gamma\left[\frac{p^{(i)}}{2}\right]\right\}\right\rangle, \qquad \text{(D2)}$$

where $p^{(i)}$ denotes the number of distinct parameters that change under model $M^{(i)}$, $i = 1, 2, 3$. Clearly, $p^{(i)} = d$, $p^{(2)} = d(d+1)/2$, and $p^{(3)} = d(d+3)/2$. Then, it has been shown that

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

$$\lim_{n\to\infty} p[W_n^{(i)} \le t] = \exp(-2e^{-t}), \quad -\infty < t < \infty, \quad i = 1, 2, 3. \tag{D3}$$

Suppose that the computed value of $W_n^{(i)} = w^{(i)}$. Then, the asymptotic p value of the test is given by

$$\lim_{n\to\infty} p[|W_n^{(i)}| > |w^{(i)}|] = 1 - \exp[-2e^{-|w^{(i)}|}] + \exp[-2e^{|w^{(i)}|}], \quad i = 1, 2, 3. \tag{D4}$$

Jarušková (1997) noted that distribution in (D3) is conservative because of the slow convergence rate.

REFERENCES

Akaike, H., 1973: Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory,* B. N. Petrov and F. Csaki, Eds., Academiai Kiado, 267–281.

Alexandersson, H., 1986: A homogeneity test applied to precipitation data. *J. Climatol.,* **6,** 661–675.

——, and A. Moberg, 1997: Homogenization of Swedish temperature data. Part I: Homogeneity test for linear trends. *Int. J. Climatol.,* **17,** 25–34.

Alley, R. B., and Coauthors, 2002: *Abrupt Climate Changes: Inevitable Surprises*. National Academy Press, 221 pp.

Angell, J. K., 1986: Annual and seasonal global temperature changes in the troposphere and low stratosphere, 1960–85. *Mon. Wea. Rev.,* **114,** 1922–1930.

——, 1999: Comparison of surface and tropospheric temperature trends estimated from a 63-station radiosonde network, 1958–1998. *Geophys. Res. Lett.,* **26,** 2761–2764.

——, 2003: Effect of exclusion of anomalous tropical stations on temperature trends from a 63-station radiosonde network, and comparison with other analyses. *J. Climate,* **16,** 2288–2295.

——, cited 2012: Global, hemispheric, and zonal temperature deviations derived from radiosonde records. *Trends Online: A Compendium of Data on Global Change,* Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN, doi:10.3334/CDIAC/cli.005.

Beaulieu, C., J. Chen, and J. L. Sarmiento, 2012: Change-point analysis as a tool to detect abrupt climate variations. *Philos. Trans. Roy. Soc. London,* **370A,** 1228–1249.

Berliner, L. M., C. Wikle, and N. Cressie, 2000: Long-lead prediction of Pacific SSTs via Bayesian dynamic modeling. *J. Climate,* **13,** 3953–3968.

Borovkov, A. A., 1999: Asymptotically optimal solutions in the change-point problem. *Theory Probab. Its Appl.* (English translation), **43,** 539–561.

Briggs, W. M., 2008: On the changes in the number and intensity of North Atlantic tropical cyclones. *J. Climate,* **21,** 1387–1402.

Cagnazzo, C., C. Claud, and S. Hare, 2006: Aspects of stratospheric long-term changes induced by ozone depletion. *Climate Dyn.,* **27,** 101–111.

Chen, J., and A. K. Gupta, 2001: On change point detection and estimation. *Commun. Stat.-Simul. Comput.,* **30,** 665–697.

Cobb, G. W., 1978: The problem of the Nile: Conditional solution to a change-point problem. *Biometrika,* **65,** 243–251.

Comiso, J. C., 2003: Warming trends in the Arctic from clear sky satellite observations. *J. Climate,* **16,** 3498–3510.

Compagnucci, R. H., M. Salles, and P. Canziani, 2001: The spatial and temporal behaviour of the lower stratospheric temperature over the Southern Hemisphere: The MSU view. Part I: Data, methodology and temporal behaviour. *Int. J. Climatol.,* **21,** 419–437.

Csörgö, M., and L. Horváth, 1997: *Limit Theorems in Change-Point Analysis.* J. Wiley and Sons, 414 pp.

Davies, R. B., 1980: Algorithm AS 155: The distribution of a linear combination of χ^2 random variables. *Appl. Stat.,* **29,** 323–333.

DeGaetano, A. T., 2006: Attributes of several methods for detecting discontinuities in temperature series: Prospects for a hybrid homogenization procedure. *J. Climate,* **9,** 1646–1660.

Ducré-Robitaille, J., L. Vincent, and G. Boulet, 2003: Comparison of techniques for detection of discontinuities in temperature series. *Int. J. Climatol.,* **23,** 1087–1101.

Easterling, D. R., and T. Peterson, 1995: A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol.,* **15,** 369–377.

Fearly, R., and J. Sweeny, 2005: Detection of a possible change point in atmospheric variability in the North Atlantic and its effect on Scandinavian glacier mass balance. *Int. J. Climatol.,* **25,** 1819–1833.

Fearnhead, P., 2005: Exact Bayesian curve fitting and signal segmentation. *IEEE Trans. Signal Process.,* **53,** 2160–2166.

——, 2006: Exact and efficient Bayesian inference for multiple change-point problems. *Stat. Comput.,* **16,** 203–213.

Feng, S., Y. Fu, and Q. Xiao, 2012: Trends in the global tropopause thickness revealed by radiosondes. *Geophys. Res. Lett.,* **39,** L20706, doi:10.1029/2012GL053460.

Field, C., 1993: Tail areas of linear combinations of central chi-squares and non-central chi-squares. *J. Statist. Comput. Simul.,* **45,** 243–248.

Forster, P. M., G. Bodeker, R. Schofield, S. Solomon, and D. Thompson, 2007: Effects of ozone cooling in the tropical lower stratosphere and upper troposphere. *Geophys. Res. Lett.,* **34,** L23813, doi:10.1029/2007GL031994.

Fotopoulos, S. B., and V. K. Jandhyala, 2001: Maximum likelihood estimation of a change-point for exponentially distributed random variables. *Stat. Probab. Lett.,* **51,** 423–429.

——, ——, and E. Khapalova, 2010: Exact asymptotic distribution of the change-point MLE for change in the mean of Gaussian sequences. *Ann. Appl. Stat.,* **4,** 1081–1104.

——, ——, and ——, 2011: Change-point MLE in the rate of exponential sequences with application to Indonesian seismological data. *J. Stat. Plann. Inference,* **141,** 220–234.

Gaffen, D. J., 1994: Temporal inhomogeneities in radiosonde temperature records. *J. Geophys. Res.,* **99,** 3667–3676.

——, M. Sargent, R. Habermann, and J. Lanzante, 2000: Sensitivity of tropospheric and stratospheric temperature trends to radiosonde data quality. *J. Climate,* **13,** 1776–1796.

Henze, N., and B. Zirkler, 1990: A class of invariant and consistent tests for multivariate normality. *Comm. Stat.: Theory Methods,* **19,** 3595–3617.

Hinkley, D. V., 1972: Time ordered classification. *Biometrika,* **59,** 509–523.

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

Hochberg, Y., and A. Tamhane, 1987: *Multiple Comparison Procedures.* J. Wiley and Sons, 450 pp.

Horváth, L., P. Kokoszka, and J. Steinebach, 1999: Testing for changes in multivariate dependent observations with an application to temperature changes. *J. Multivariate Anal.,* **68,** 96–119.

Imhof, J. P., 1961: Computing the distribution of a quadratic form in normal variables. *Biometrika,* **48,** 419–426.

Jandhyala, V. K., and S. B. Fotopoulos, 1999: Capturing the distributional behavior of the maximum likelihood estimator of a change-point. *Biometrika,* **86,** 129–140.

——, P. Liu, and S. B. Fotopoulos, 2009: River stream flows in the northern Québec Labrador region: A multivariate change point analysis via maximum likelihood. *Water Resour. Res.,* **45,** W02408, doi:10.1029/2007WR006499.

Jarušková, D., 1996: Change-point detection in meteorological measurement. *Mon. Wea. Rev.,* **124,** 1535–1543.

——, 2010: Asymptotic behavior of a test statistic for detection of change in mean of vectors. *J. Stat. Plann. Inference,* **140,** 616–625.

Karcher, M., R. Gerdes, F. Kauker, and C. Koberle, 2003: Arctic warming—Evolution and spreading of the 1990s warm event in the Nordic seas and the Arctic Ocean. *J. Geophys. Res.,* **108,** 3034, doi:10.1029/2001JC001265.

Karl, T. R., S. J. Hassol, C. D. Miller, and W. L. Murray, Eds., 2006: *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences.* U.S. Climate Change Science Program and Subcommittee on Global Change Research, 164 pp. [Available online at http://downloads.globalchange.gov/sap/sap1-1/sap1-1-final-all.pdf.]

Kuonen, D., 1999: Saddlepoint approximations for distributions of quadratic forms in normal variables. *Biometrika,* **86,** 929–935.

Lanzante, J., A. Stephen, and D. Siedel, 2003: Temporal homogenization of monthly radiosonde temperature data. *J. Climate,* **16,** 224–240.

Lau, K. M., and H. Wu, 2007: Detecting trends in tropical rainfall characteristics, 1979–2003. *Int. J. Climatol.,* **27,** 979–988.

Lu, Z., 2006: The numerical evaluation of the probability density function of a quadratic form in normal variables. *Comput. Stat. Data Anal.,* **51,** 1986–1996.

Lund, R., and J. Reeves, 2002: Detection of undocumented change points: A revision of the two-phase regression model. *J. Climate,* **15,** 2547–2554.

Mardia, K. V., 1970: Measures of multivariate skewness and kurtosis with applications. *Biometrika,* **57,** 519–530.

McShane, B. B., and A. J. Wyner, 2011: A statistical analysis of multiple temperature proxies: Are reconstructions of surface temperatures over the last 1000 years reliable? *Ann. Appl. Stat.,* **5,** 5–44.

Perreault, L., M. Haché, M. Slivitzky, and B. Bobée, 1999: Detection of changes in precipitation and runoff over eastern Canada and U.S. using Bayesian approach. *Stochastic Environ. Res. Risk Assess.,* **13,** 201–216.

——, J. Bernier, B. Bobée, and É. Parent, 2000a: Bayesian change point analysis in hydrometeorological time series 1, part 1. *J. Hydrol.,* **235,** 221–241.

——, ——, ——, and ——, 2000b: Bayesian change point analysis in hydrometeorological time series 1, part 2. *J. Hydrol.,* **235,** 242–263.

——, É. Parent, J. Bernier, B. Bobée, and M. Slivitzky, 2000c: Retrospective multivariate Bayesian change-point analysis: A simultaneous single change in the mean of several hydrological sequences. *Stochastic Environ. Res. Risk Assess.,* **14,** 243–261.

Peterson, T. C., and Coauthors, 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.,* **18,** 1493–1517.

Ramaswamy, V., and Coauthors, 2001: Stratospheric temperature trends: Observations and model simulations. *Rev. Geophys.,* **39,** 71–122.

Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 589–662.

Randel, W. J., and F. Wu, 1999: Cooling of the Arctic and Antarctic polar stratospheres due to ozone depletion. *J. Climate,* **12,** 1467–1479.

——, and Coauthors, 2009: An update of observed stratospheric temperature trends. *J. Geophys. Res.,* **114,** D02107, doi:10.1029/2008JD010421.

Reeves, J., J. Chen, X. Wang, R. Lund, and Q. Lu, 2007: A review and comparison of changepoint detection techniques for climate data. *J. Appl. Meteor. Climatol.,* **46,** 900–915.

Rodionov, S. N., 2004: A sequential algorithm for testing climate shifts. *Geophys. Res. Lett.,* **31,** L09204, doi:10.1029/2004GL019448.

Ruggieri, E., 2012: A Bayesian approach to detecting change points in climatic records. *Int. J. Climatol.,* **33,** 520–528, doi:10.1002/joc.3447.

Santer, B. D., and Coauthors, 2003: Contributions of anthropogenic and natural forcing to recent tropopause height changes. *Science,* **301,** 479–483.

Schleip, C., A. Menzel, and V. Dose, 2009: Bayesian analysis of changes in radiosonde atmospheric temperature. *Int. J. Climatol.,* **29,** 629–641.

Seidel, D. J., and J. R. Lanzante, 2004: An assessment of three alternatives to linear trends for characterizing global atmospheric temperature changes. *J. Geophys. Res.,* **109,** D14108, doi:10.1029/2003JD004414.

——, and W. J. Randel, 2006: Variability and trends in the global tropopause estimated from radiosonde data. *J. Geophys. Res.,* **111,** D21101, doi:10.1029/2006JD007363.

——, N. P. Gillett, J. R. Lanzante, K. P. Shine, and P. W. Thorne, 2011: Stratospheric temperature trends: Our evolving understanding. *WIREs. Climatic Change,* **2,** 591–616.

Seidou, O., and T. Ouarda, 2007: Recursion-based multiple change point detection in multiple linear regression and application to river streamflows. *Water Resour. Res.,* **43,** W07404, doi:10.1029/2006WR005021.

——, J. Asselin, and T. Ouarda, 2007: Bayesian multivariate linear regression with application to change point models in hydrometeorological variables. *Water Resour. Res.,* **43,** W08401, doi:10.1029/2005WR004835.

Solomon, S., R. Portmann, and D. Thompson, 2007: Contrasts between Antarctic and Arctic ozone depletion. *Proc. Natl. Acad. Sci. USA,* **104,** 445–449.

Son, S.-W., N. F. Tandon, and L. M. Polvani, 2011: The fine-scale structure of the global tropopause derived from COSMIC GPS radio occultation measurements. *J. Geophys. Res.,* **116,** D20113, doi:10.1029/2011JD016030.

Son, Y. S., and S. Kim, 2005: Bayesian single change point detection in a sequence of multivariate normal observations. *Statistics,* **39,** 373–387.

Steinbrecht, W., B. Hassler, H. Claude, P. Winkler, and R. Stolarski, 2003: Global distribution of total ozone and lower stratospheric temperature variations. *Atmos. Chem. Phys.,* **3,** 1421–1438.

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC

Thorne, P. W., and R. S. Vose, 2010: Reanalyses suitable for characterizing long-term trends: Are they really achievable? *Bull. Amer. Meteor. Soc.,* **91,** 353–361.

——, J. R. Lanzante, T. C. Peterson, D. J. Seidel, and K. P. Shine, 2010: Tropospheric temperature trends: History of an ongoing controversy. *WIREs. Climatic Change,* **2,** 66–88, doi:10.1002/wcc.80.

Vincent, L. A., 1998: A technique for the identification of inhomogeneities in Canadian temperature series. *J. Climate,* **11,** 1094–1104.

Wang, X. L., Q. Wen, and Y. Wu, 2007: Penalized maximal *t* test for detecting undocumented mean change in climate data series. *J. Appl. Meteor. Climatol.,* **46,** 916–931.

Zhao, X., and P. Chu, 2006: Bayesian multiple change-point analysis of hurricane activity in the eastern North Pacific: A Markov chain Monte Carlo approach. *J. Climate,* **19,** 564–578.

Zou, C.-Z., and W. Wang, 2010: Stability of the MSU-derived atmospheric temperature trend. *J. Atmos. Oceanic Technol.,* **27,** 1960–1971.

Unauthenticated | Downloaded 02/23/24 09:27 AM UTC