Journal of Environmental Statistics
February 2013, Volume 4, Issue 3. http://www.jenvstat.org

Change-Point Analysis of Annual Mean Precipitation for Northern, Tropical and Southern Latitudes of the Globe in the Past Century

Elena A Khapalova University of Michigan-Flint
Venkata K. Jandhyala Washington State University
Stergios B. Fotopoulos Washington State University

Abstract

Applying methods of change-point analysis, a statistical study of the annual mean precipitation from northern, tropical and southern latitudes of the globe is carried out based upon data for the past 100 years. The change-point analysis assumes multivariate Gaussianity for the time-series data, thus allowing spatial correlations between the three latitudes. Recently developed change detection and change-point estimation methods for multivariate Gaussian data are applied to analyze the data. A simulation study is also carried out in a bivariate set-up to investigate the robustness of the estimation methodology for deviation from Gaussianity as well as its sensitivity to estimation of the model parameters. Data analysis identifies a significant change in the precipitation from northern and southern latitudes subsequent to the year 1944, whereas no change was identified in the data from tropical latitude. The discussion suggests that the change found in the data from northern and southern latitudes may not have been entirely due to gauge changes introduced in and around 1950.

Keywords: mean precipitation, change-point analysis, likelihood ratio test, multivariate Gaussian distribution.

1. Introduction

Numerous climatic indicators are used in the literature for tracking changes in the climate, and some of the commonly used ones are: droughts, moisture surpluses, surface temperatures, precipitation, to name a few. Of these, surface air temperature and precipitation have been monitored the longest in time. Hence, recorded data on these variables can be utilized for detecting possible changes over time (e.g. Karl, Knight, Easterling, and Quayle (1996)). Both temperature and precipitation directly affect human lives and activities, and the environment.

Thus, changes in these climatological variables are of immediate concern for scientists and people alike.

Scientists utilize climate models for purposes of forecasting as well as monitoring climate changes (e.g. Karl *et al.* (1996), Smith and Lazo (2001), Evans (2001)). Such scientific monitoring techniques are important as countries around the globe adapt to projections regarding future climatic scenarios. From the perspective of monitoring precipitation over time, significant changes in precipitation may cause flooding or drought, and this in turn affects agriculture. Changes in precipitation also affect crop irrigation, soil erosion, and the type of crops that may grow in a given area. While some countries may benefit, many others will suffer decreased yields. For these reasons, those who model crop yield and production use long term precipitation averages as one of the main input variables in their models (Hubbard and Flores-Mendoza (1995), Perarnaud, Seguin, Malezieux, Deque, and Loustau (2005), Tao, Yokozawa, Hayashi, and Lin (2005), Evans (2009)).

In this article, the focus is on identifying changes in mean annual precipitation through recorded data for northern, tropical and southern latitudes of the globe. Past and even recent studies on precipitation changes have generally focused their analyses on regional or country specific level data (Gutman, Hosking, and Wallis (1993), Hurrel (1995), Ho, Lee, Ahn, and Lee (2003), Kripalani, Kulkarni, Sabade, and Khandekar (2003), Koning and Franses (2005), Buda, Tong, Guoyu, and Zhenghong (2007), Nastos and Zaferos (2007), Goubanova and Li (2007), Kioutsioukis, Melas, and Zeferos (2009), Choi, Kim, and Byun (2009)). However, precipitation levels and their effects do not limit themselves to artificial country or regional boundaries. Taking a global point of view, Giannini, Biasutti, Held, and Sobel (2008) point out that it is global climate that primarily influences changes in the African climate, and that African rainfall is affected not just by local natural events and human activities but through global-scale mechanisms. The researchers point out that one requires both global as well as local analyses of climate variables to truly comprehend the observed changes in the environment, what is driving those changes, and how current and future changes should be managed. Thus, taking global climatic patterns into account is important for a better understanding of how climate change affects not just smaller regions, but larger areas such as an entire hemisphere, or even the entire globe. With this objective in mind, this article analyzes recorded mean annual precipitations within each of northern, tropical and southern latitudes for the past century in order to identify instances of possible changes that may have occurred in the past 100 years.

Various approaches are considered in the literature for identifying changes in climatic variables. First, climatologists and atmospheric scientists base their analysis and conclusions on a careful understanding of the scientific phenomena, and how climatic variables may change from a scientific perspective. One, however, awaits empirical evidence in favor of such postulates and often such empirical evidence comes from proper modeling of recorded data. In this regard, Reeves, Chen, Wang, Lund, and Lu (2007) gives an overview of the most common change-point detection techniques used in climate literature including the standard normal homogeneity test, nonparametric SNH test, Akaike's information criteria, and Sawa's Bayes criteria. For example, Gonzalez-Rouco, Jimenez, Quesada, and Valero (2001) use the standard normal homogeneity test, while Hegerl and North (1997) compare the three statistically optimal approaches of weighted average, fingerprinting, and filtering for detecting climate change. A Bayesian approach to fingerprinting technique has been adapted by Berliner, Levine, and Shea (2000).

Overall, change-point methodology has recently become an important tool for identifying sustained changes in climatic variables (Lund and Reeves (2002), Briggs (2008), Zhao and Chu (2006), DeGaetano (2006), Easterling and Peterson (1995), Fearly and Sweeny (2005), Rodionov (2004), Jaruskova (1996), Lau and Wu (2007), Wang, Wen, and Wu (2007), Jandhyala, Liu, and Fotopoulos (2009), Fotopoulos, Jandhyala, and Khapalova (2010)). Change-point methodology begins with the detection part where one applies change detection statistics to establish evidence for the presence of one or more changes in a given series. When the presence of a change-point is confirmed, then one applies an estimation technique in order to estimate the location of the unknown change-point. The purpose of this article is to pursue changes in global precipitation in three zones by implementing change-point methodology for Gaussian sequences in a multivariate framework. Most recently, Fotopoulos *et al.* (2010) derived an exact computational procedure for estimating an unknown change-point in a multivariate series. The goal of the present article is to take advantage of this computational procedure for a comprehensive understanding of the time-periods at which changes in global mean precipitation may have occurred in the past century. The multivariate approach we take in this article is quite recent and accounts for spatial relationships that might be present between the three global zones.

The paper is organized as follows. Section 2 describes the global precipitation data that we analyze in the paper. Section 3 summarizes change-point methodology including both detection as well as the maximum likelihood estimation (mle) of an unknown change-point. Since the methodology assumes the data to be Gaussian and that the data series is time independent, one may be concerned about the validity of this assumption for real data. A proper understanding of the sensitivity of the methodology to deviations from normality could alleviate such concerns if the method is found to be quite robust. In this regard, we perform in Section 4 extensive simulations for the bivariate case, going beyond the univariate sensitivity studies reported in Fotopoulos *et al.* (2010), as well as Jandhyala *et al.* (2009). Section 5 consists of change-point analysis of the global precipitation data. In Section 6, we evaluate change-point models from the viewpoint of model selection through the AIC criterion and Section 7 concludes the article with discussion and concluding remarks.

2. Data

The data on annual mean precipitation in mm per day for three latitude bands covering 30%, 40%, and 30% of the global area, which is the central focus of this article, is originally reported by Mitchell, Carter, Jones, Hulme, and New (2004) and covers the period 1901-2000. The data is a part of a much larger comprehensive set of high-resolution grid data on monthly climatic conditions for Europe and the Globe. The dataset covers the major climate zones: tropical, temperate, and arctic, with the temperate and arctic zones being grouped together. The Northern Latitudes (NL) band (24° N to 90° N) covers the regions of North America, Europe, and Asia. The Low Latitudes (LL) band (24° S to 24° N) covers the regions of Africa, Asia, Australia, Central and South America. Finally, the Southern Latitudes (SL) band (24° S to 90° S) covers the regions of South America, Africa, Antarctica, and Australia. Figure 1 below displays the time series plot of the data on annual mean precipitation in mm per day for the years 1901-2000 for the NL, Low and SL bands of the globe.

Before proceeding any further with a discussion on the data and its analysis, we address the issue of whether the three latitudes (NL, LL and SL) considered above are sensible for studying the dynamic behavior of global precipitation. From the view-point of weather patterns, the LL latitude represents the tropical zone, whereas the NL latitude represents the northern temperate and arctic zone, and the SL latitude represents the southern temperate zone. In this sense, the grouping seems to represent three distinct weather zones at a global level. Moreover, there is additional discussion in the literature in support of considering these three zones. The Hadley cell effect, which speaks of air circulating from the tropics (LL zone) to regions approximately 30° north and south latitudes (NL and SL zones), where the air masses sink seems to have considerable effect on the dynamic behavior of global climate. While the Hadley cell effect itself is not new in the literature, recently, Stevens (2011) discussed the influence of Hadley cell effect on global climate. Zhou, Xu, Sud, and Betts (2011) showed more explicitly that global precipitation is affected by the Hadley cell effect phenomenon. In light of these observations, it seems natural to consider northern (NL), tropical (LL), and southern (LL) latitudes (as suggested by the Hadley cell effect) for studying the dynamic behavior of global precipitation.

The precipitation data in Figure 1 used in this article has not been corrected for gauge bias either due to under-catch of solid precipitation or periodic instrument changes. It is well known that correction for global datasets is extremely time intensive and requires data from individual stations that is not readily available. Since only land data is used, lack of sufficient data on oceans is not a concern in this case. Multiple papers (e.g. Karl, Quayle, and Groisman (1993), Karl *et al.* (1996), New, Todd, Hulme, and Jones (2001), Hulme (1995), Gonzalez-Rouco *et al.* (2001), Hegerl, Karl, Allen, Bindoff, Gillett, Karoly, Zhang, and Zwiers (2006)) have mentioned various issues that may arise with precipitation data collected over the past century. One may refer to Figure 1 of Karl *et al.* (1993) for gauge changes and precipitation record discontinuities in different countries over the past 100 years.

When a change-point in a precipitation dataset is detected, one must then enquire whether the detected change is in the precipitation itself or whether it is due to changes in observational practices. Many articles have been written on detecting and correcting for artificial mean shifts in climate data, which may occur due to instrument or station location change, gauge change, calibration issues, and other reasons as well. Of special concern are “undocumented shifts”that may have occurred when a station has made changes without documenting them. Multiple approaches have been taken to locate such shifts, and then to correct for them as much as possible with given information; this is generally referred to as homogenization of the climate record, (see Wang (2003), Wang (2008a), Wang (2008b), Menne and Williams (2005), Sherwood (2007), Gonzalez-Rouco *et al.* (2001), Hegerl *et al.* (2006)).

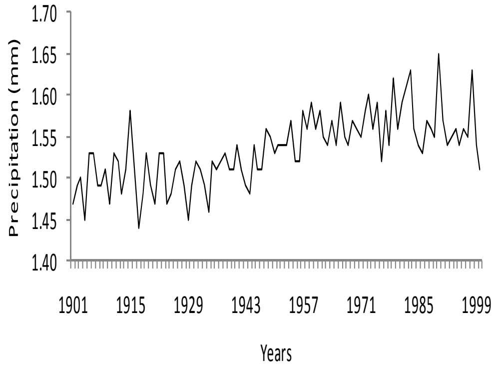

Figure 1(a) — Northern Latitude (NL)

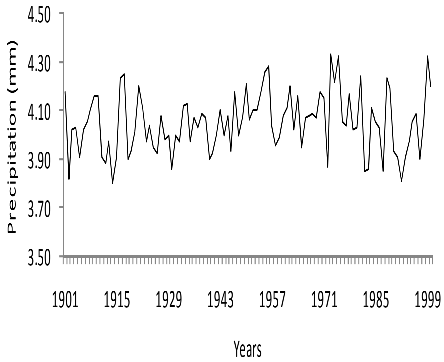

Figure 1(b) — Low Latitude (LL)

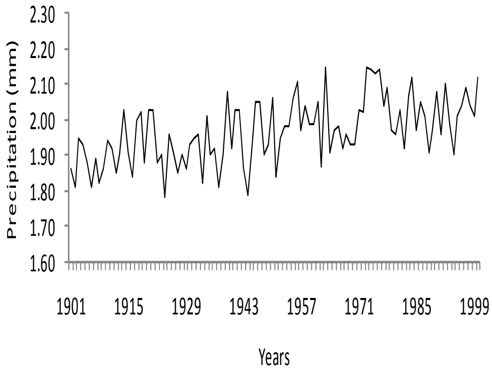

Figure 1(c) — Southern Latitude (SL)

Figure 1. Annual mean precipitation data for NL, LL, and SL zones of the globe for the years 1901-2000. (a) Precipitation for NL zone. (b) Precipitation for LL zone. (c) Precipitation for SL zone.

3. Methodology

The statistical methodology we adapt to analyze the data in Figure 1 is referred to as change-point analysis and consists of two main parts, namely, detection of change-points and then estimation of change-points including confidence interval estimation. For a comprehensive overview of the theory on change-point detection and estimation, one may see Csörgö and Horváth (1997) and Chen and Gupta (2000). Here, we only present details of the methodology that is specific to the analysis of the precipitation data. For purposes of analyzing the data, we assume that the data can be modeled by the multivariate Gaussian family, and that the data are independent over time. It is also assumed that changes may occur only in the mean vector with the covariance structure between the three zones remaining stable over time. Here, we would like to comment regarding the assumption of independence over time. If one were to look for serial correlations (based on ACF and PACF plots) directly in each of the three data series, then one does indeed identify autoregressive correlations up to lag 8, particularly in the NL and SL series. However, these serial correlations disappear once the change-point model is incorporated. Thus, we believe that the time dependence in the data series is merely a consequence of not accommodating change-point in the mean of the data series, and otherwise it is not inherent to the data series itself. In fact, subsequent residual analysis shows that all the three assumptions including the assumption of independence over time are indeed appropriate for the data in this study.

We shall first present the method of detecting an unknown change-point in the mean vector of a multivariate Gaussian series of dimension d (clearly, $d = 3$ for the data in Figure 1). The likelihood ratio statistic we apply here has been developed in Csörgö and Horváth (1997). Let the observed d-dimensional data of size n be denoted by $Y_1, Y_2, ..., Y_n, n \geq 1$, wherein we let each observation follow the multivariate Gaussian distribution with mean vector μ, and covariance matrix Σ. The underlying parameter (μ, Σ) is assumed to begin with an initial value of (μ_1, Σ) and subsequently change to (μ_2, Σ), at some unknown index point $\tau_n \in \{1, 2, ..., n-1\}$. Asymptotic distribution theory of the generalized likelihood ratio statistic for testing the presence of an unknown change-point τ_n has been derived in Csörgö and Horváth (1997). Specifically, the corresponding twice log-likelihood ratio statistic is:

$$U_n = max_{1 \leq t \leq n-1} nlog \left(\left| \hat{\Sigma}_n \right| / \left| \hat{\Sigma}_t \right| \right) \tag{3.1}$$

where $\hat{\Sigma}_t = n^{-1} \left\{ \sum_{i=1}^{t} \left(Y_i - \hat{\mu}_{1,t}\right) \left(Y_i - \hat{\mu}_{1,t}\right)^T + \sum_{i=t+1}^{n} \left(Y_i - \hat{\mu}_{2,t}\right) \left(Y_i - \hat{\mu}_{2,t}\right)^T \right\}$,
$\hat{\mu}_{1,t} = t^{-1} \sum_{i=1}^{t} Y_i,\ \hat{\mu}_{2,t} = (n-t)^{-1} \sum_{i=t+1}^{n} Y_i,\ t = 1, 2, ..., n$. Letting

$$W_n = (2log\,log\,n\,U_n)^{1/2} - \left(2log\,log\,n + \frac{d}{2} log\,log\,log\,n - log\Gamma\left(\frac{d}{2}\right) \right)$$

enables one to find critical values of the statistic in (3.1) via the limiting distribution of W_n:

$$lim_{n\to\infty}P\left[W_n \leq t\right] = exp\left(-2e^{-t}\right). \tag{3.2}$$

When the above test results in significance, the mle $\hat{\tau}_n$ of the unknown change-point τ_n is obtained as the position at which U_n attains its maximum.

Asymptotic distribution of the mle $\hat{\tau}_n$ can be pursued through the centered estimator given by $\xi_n = \hat{\tau}_n - \tau_n$. The limiting distribution of ξ_n denoted by ξ_∞ has been recently derived by Fotopoulos *et al.* (2010) in a form that is fully computable. We present here only the basic result, and refer to their paper for details regarding its computational aspects. The required result for the multivariate case is presented in Section 3.2 in Fotopoulos *et al.* (2010). Assuming the model parameters μ_1, μ_2, Σ to be known, it has been shown that:

$$P\left(\xi_\infty = k\right) = \begin{cases} \left(1-\|G_+\|\right)\left(q_{|k|} - \|G_+\|\,\tilde{q}_{|k|}\right) & k = \pm 1, \pm 2, ... \\ \left(1-\|G_+\|\right)^2 & k = 0. \end{cases} \tag{3.3}$$

where $(1 - \|G_+\| = exp\left\{-\sum_{j=1}^{\infty}\frac{1}{j}\bar{\Phi}\left(\eta\sqrt{j}/2\right)\right\}$, $\eta^2 = \left(\mu_1 - \mu_2\right)^T \Sigma^{-1}\left(\mu_1 - \mu_2\right)$, and $\bar{\Phi}\left(\cdot\right)$ is the survival function of the univariate standard normal distribution. Also, $q_k = E\left\{I\left(T_1^- > k\right)\right\}$, $\tilde{q}_k = E\left\{e^{-S_k}I\left(T_1^- > k\right)\right\}$, k = 1, 2,..., and $q_0 = \tilde{q}_0 = 1$ where T_1^- represents the first time that a random walk with negative drift becomes negative. Here, assumption of model parameters being known can be viewed as a limitation, since in practice, the model parameters have to be estimated from the observed data. However, Hinkley (1972) has shown that the distribution in (3.3) remains unchanged even when the unknown parameters are replaced by their estimators. The practical applicability of this result, however, can be cross-validated through simulations, which we do so in the next section.

For estimating the unknown change-point, one may alternatively implement the conditional solution of Cobb (1978), which is otherwise known as conditional maximum likelihood estimator (cmle), and we pursue this methodology also in our analysis. This solution is attractive mainly because it can be viewed as a Bayesian solution to the problem of estimating an unknown change-point. If δ denotes the number of data points to be considered on either side of $\hat{\tau}_n$, with the belief that the true change-point lies within the δ neighborhood of $\hat{\tau}_n$, then Cobb's conditional solution is given by

$$P\left(\xi_n = l|Y_{\hat{\tau}_n-\delta+1},...,Y_{\hat{\tau}_n+\delta}\right) \cong p_n\left(Y;\hat{\tau}_n+l\right)/\sum_{d=-\delta}^{\delta}p_n\left(Y;\hat{\tau}_n+l\right), l \in \{-\delta,...,\delta\} \tag{3.4}$$

where $p_n(Y;t)$ denotes the likelihood function of the data series with change-point at t. The method of choosing δ is detailed in Cobb (1978).

4. Simulations

Here, the main goal is to investigate empirically the robustness of the mle and Cobb's conditional solution to deviations from Gaussianity. Also, another important goal of this section is to study empirically the distributional equivalence of the change-point estimate under known and estimated parameter situations. The simulation study is carried out for various sample sizes, amounts of changes as well as various locations of the change-point. The univariate case of this simulation study has been carried out in Jandhyala *et al.* (2009) as well as Fotopoulos *et al.* (2010) and the results for deviations from Gausianity were quite satisfactory. It is important to carry out a similar study for higher dimensions because the multivariate case involves estimation of many more unknown parameters. Thus, the simulation study in this article addresses the bivariate case, and if the results are satisfactory in the bivariate case also, one may at that point conclude that the same would continue to hold for three and higher dimensions. Noting that simulation results for deviations from independence in the univariate case were not as good (see Jandhyala *et al.* (2009)), we limit the simulations here to departures from Gaussianity only.

We formulate departures from bivariate Gaussianity mainly by modeling the error structure through the standardized bivariate t(ν)-distribution with degrees of freedom chosen to be $\nu = 5, 10, 20$. As is well-known, the bivariate t(ν)-distribution is defined by scaling the bivariate standard Gaussian distribution in the denominator with the square root of an independent univariate chi-square distribution with ν degrees of freedom divided by ν. The combinations of n and τ_n (henceforth depicted as τ) we chose in the simulations are: $n = 50, \tau = 25; n = 100, \tau = 25; n = 100, \tau = 50; n = 200, \tau = 50; n = 200, \tau = 100$. Each of these combinations are run for values of: $\eta = 1.0, 1.5, 2.0, 2.5$. Based upon 100,000 simulations for each case, we compute bias (Bias) and square root of the mean square error (RMSE) for the empirical distributions of change-point mle and the conditional solution of Cobb (1978). The results for Bias and RMSE for known η are presented in Table 1 and for the case of estimated η in Table 2.

Clearly, the results for Bias are extremely satisfactory for both known and unknown cases uniformly throughout the various parameter choices. The results for RMSE are also extremely good in the known case (Table 1) throughout the parameters in the sense that the empirical values are quite close to the corresponding theoretical values even when sample size is 50, and amount of change η is small. The same is true for the estimated case (Table 2) also, except when $\eta = 1.0$, in which case, the closeness is marginally good. As in the univariate case (see Figure 1 in Fotopoulos *et al.* (2010)), the change-point mle has smaller RMSE values compared to Cobb's cmle uniformly under both known and unknown cases.

			Bias						Square root of MSE					
				n = 200		n = 100		n = 50		n = 200		n = 100		n = 50
			Theor.	τ = 150	τ = 100	τ = 75	τ = 50	τ = 25	Theor.	τ = 150	τ = 100	τ = 75	τ = 50	τ = 25
gauss	η = 1.0	mle	0.000	0.014	-0.024	-0.016	-0.011	0.003	5.156	5.167	5.044	4.941	5.088	4.724
		cmle	0.000	0.022	-0.018	-0.046	0.008	-0.021		6.200	6.140	5.950	6.141	5.688
	η = 1.5	mle	0.000	-0.009	0.015	-0.011	-0.003	0.004	2.267	2.241	2.283	2.257	2.214	2.234
		cmle	0.000	-0.007	0.012	-0.005	-0.004	0.000		2.738	2.742	2.728	2.709	2.710
	η = 2.0	mle	0.000	0.009	0.001	-0.008	0.015	0.010	1.246	1.248	1.229	1.250	1.245	1.243
		cmle	0.000	0.006	0.003	-0.006	0.011	0.008		1.515	1.507	1.514	1.519	1.505
	η = 2.5	mle	0.000	0.000	-0.006	0.007	0.006	0.002	0.769	0.757	0.761	0.764	0.768	0.766
		cmle	0.000	0.001	-0.003	0.004	0.006	-0.002		0.933	0.930	0.935	0.935	0.936
t (5)	η = 1.0	mle	0.000	-0.007	-0.042	-0.065	0.062	0.014	5.156	5.079	5.126	4.811	5.125	4.651
		cmle	0.000	0.000	-0.031	-0.090	0.041	0.016		6.146	6.173	5.899	6.198	5.661
	η = 1.5	mle	0.000	0.006	-0.007	-0.016	0.001	-0.020	2.267	2.272	2.313	2.239	2.284	2.237
		cmle	0.000	0.000	-0.003	-0.008	0.002	-0.020		2.752	2.745	2.744	2.760	2.743
	η = 2.0	mle	0.000	-0.002	-0.001	-0.004	0.001	0.005	1.246	1.292	1.356	1.274	1.298	1.262
		cmle	0.000	-0.005	0.001	0.001	0.006	0.008		1.549	1.548	1.535	1.548	1.533
	η = 2.5	mle	0.000	0.000	-0.001	0.000	-0.004	-0.007	0.769	0.813	0.818	0.811	0.815	0.813
		cmle	0.000	-0.001	0.001	0.000	-0.001	-0.005		0.966	0.969	0.965	0.971	0.964
t (10)	η = 1.0	mle	0.000	0.012	-0.002	-0.046	-0.020	-0.013	5.156	5.063	5.145	4.883	5.071	4.728
		cmle	0.000	0.028	0.005	-0.088	-0.020	-0.009		6.151	6.215	5.930	6.145	5.689
	η = 1.5	mle	0.000	-0.011	0.010	-0.009	0.004	0.006	2.267	2.223	2.246	2.241	2.251	2.218
		cmle	0.000	-0.004	0.010	-0.010	0.006	0.001		2.722	2.736	2.722	2.736	2.704
	η = 2.0	mle	0.000	0.001	-0.005	0.006	0.004	-0.001	1.246	1.255	1.228	1.245	1.232	1.217
		cmle	0.000	0.000	-0.003	0.006	0.002	-0.002		1.526	1.501	1.516	1.510	1.497
	η = 2.5	mle	0.000	-0.003	0.001	0.000	0.000	0.001	0.769	0.771	0.765	0.793	0.783	0.767
		cmle	0.000	-0.003	-0.001	0.000	0.001	0.003		0.941	0.934	0.951	0.944	0.937
t (20)	η = 1.0	mle	0.000	-0.029	-0.003	-0.057	0.019	0.036	5.156	5.088	5.055	4.926	4.997	4.743
		cmle	0.000	-0.010	0.007	-0.094	0.025	0.015		6.156	6.165	5.945	6.105	5.686
	η = 1.5	mle	0.000	0.000	0.003	0.002	0.007	0.011	2.267	2.240	2.231	2.249	2.241	2.215
		cmle	0.000	0.001	0.001	0.002	0.007	0.006		2.728	2.711	2.740	2.733	2.712
	η = 2.0	mle	0.000	-0.003	0.013	-0.004	0.003	-0.004	1.246	1.231	1.234	1.236	1.224	1.240
		cmle	0.000	0.002	0.009	-0.002	0.002	-0.003		1.506	1.510	1.509	1.496	1.505
	η = 2.5	mle	0.000	0.002	0.001	0.004	0.004	-0.009	0.769	0.767	0.762	0.765	0.770	0.771
		cmle	0.000	0.004	0.002	0.005	0.002	-0.007		0.937	0.934	0.938	0.934	0.938

Table 1. Bias and RMSE for the empirical distribution (based on 100000 simulations) of change-point mle and cmle for the case of η known.

			Bias						Square root of MSE					
				n = 200		n = 100		n = 50		n = 200		n = 100		n = 50
			Theor.	τ = 150	τ = 100	τ = 75	τ = 50	τ = 25	Theor.	τ = 150	τ = 100	τ = 75	τ = 50	τ = 25
gauss	η = 1.0	mle	0.000	-0.421	-0.033	-1.597	-0.001	-0.042	5.156	7.498	6.254	10.612	7.427	7.516
		cmle	0.000	-0.740	-0.027	-2.652	-0.011	-0.048		8.769	7.537	12.284	8.523	7.880
	η = 1.5	mle	0.000	-0.072	0.022	-0.193	0.002	-0.005	2.267	2.558	2.426	3.131	2.682	3.224
		cmle	0.000	-0.122	0.017	-0.310	-0.001	-0.006		3.125	2.986	3.765	3.287	3.782
	η = 2.0	mle	0.000	-0.009	0.003	-0.054	0.010	0.009	1.246	1.344	1.284	1.444	1.385	1.538
		cmle	0.000	-0.030	0.006	-0.087	0.010	0.010		1.635	1.593	1.781	1.703	1.904
	η = 2.5	mle	0.000	-0.009	-0.006	-0.017	0.007	0.001	0.769	0.790	0.783	0.840	0.812	0.901
		cmle	0.000	-0.014	-0.004	-0.028	0.005	-0.001		0.982	0.967	1.044	1.010	1.117
t (5)	η = 1.0	mle	0.000	-0.735	0.002	-2.141	-0.023	0.027	5.156	10.897	7.832	13.028	9.075	8.133
		cmle	0.000	-1.084	-0.022	-3.037	-0.040	0.017		11.808	8.821	13.932	9.765	8.280
	η = 1.5	mle	0.000	-0.080	-0.003	-0.255	-0.005	-0.007	2.267	3.504	3.068	4.524	3.302	3.871
		cmle	0.000	-0.134	-0.004	-0.412	0.002	-0.027		3.942	3.494	5.042	3.806	4.256
	η = 2.0	mle	0.000	-0.022	0.000	-0.066	0.013	0.000	1.246	1.669	1.411	2.163	1.662	1.950
		cmle	0.000	-0.044	0.001	-0.101	0.013	-0.001		1.924	1.677	2.409	1.932	2.223
	η = 2.5	mle	0.000	-0.008	0.008	-0.016	-0.006	-0.002	0.769	0.858	1.136	1.217	0.998	1.154
		cmle	0.000	-0.014	0.008	-0.033	-0.005	-0.003		1.029	1.265	1.364	1.158	1.324
t (10)	η = 1.0	mle	0.000	-0.486	-0.003	-1.875	0.004	0.002	5.156	8.053	6.361	11.765	8.038	7.845
		cmle	0.000	-0.801	0.001	-2.816	-0.035	-0.040		9.322	7.607	12.888	8.907	8.077
	η = 1.5	mle	0.000	-0.072	0.008	-0.202	0.017	0.000	2.267	2.517	2.428	3.384	2.731	3.353
		cmle	0.000	-0.128	0.014	-0.334	0.008	-0.023		3.110	2.982	4.001	3.325	3.871
	η = 2.0	mle	0.000	-0.022	-0.003	-0.037	0.001	-0.005	1.246	1.334	1.282	1.417	1.372	1.561
		cmle	0.000	-0.037	-0.002	-0.075	-0.001	-0.003		1.641	1.584	1.762	1.699	1.918
	η = 2.5	mle	0.000	-0.011	0.000	-0.017	0.003	0.000	0.769	0.808	0.786	0.850	0.854	0.890
		cmle	0.000	-0.016	-0.002	-0.029	0.001	0.004		0.990	0.968	1.047	1.044	1.111
t (20)	η = 1.0	mle	0.000	-0.456	-0.013	-1.748	-0.005	0.020	5.156	7.520	6.295	11.183	7.559	7.630
		cmle	0.000	-0.794	0.007	-2.757	-0.015	-0.002		8.868	7.571	12.581	8.548	7.941
	η = 1.5	mle	0.000	-0.068	0.003	-0.175	-0.006	0.020	2.267	2.530	2.415	3.210	2.671	3.328
		cmle	0.000	-0.124	-0.003	-0.296	-0.001	0.007		3.106	2.950	3.828	3.276	3.863
	η = 2.0	mle	0.000	-0.018	0.017	-0.051	0.006	0.002	1.246	1.316	1.282	1.437	1.333	1.542
		cmle	0.000	-0.031	0.009	-0.080	0.001	0.000		1.626	1.594	1.770	1.670	1.916
	η = 2.5	mle	0.000	-0.004	0.003	-0.008	0.003	-0.005	0.769	0.798	0.791	0.849	0.813	0.876
		cmle	0.000	-0.011	0.002	-0.023	0.001	-0.005		0.990	0.973	1.052	1.014	1.103

Table 2. Bias and RMSE for the empirical distribution (based on 100000 simulations) of change-point mle and cmle for the case of η unknown.

5. Data analysis and results

The goal of this section is to pursue the presence of change-points in the precipitation series (Figure 1) by applying the change-point methodology discussed in Section 3. To begin applying the methodology, we assume that the data series for NL, LL, and SL zones follows the 3-dimensional multivariate Gaussian distribution wherein the mean vector may have changed from an initial value of μ_1 to an alternative value of μ_2 at some unknown change-point τ, while the covariance Σ remains constant throughout the sampling period. We then apply the likelihood ratio statistic in (3.1) for detecting the presence of a change-point τ. As we apply (3.1), we perform detection for every one of the three zones at a univariate level, at a bivariate level for every pair, and then for all three zones under the 3-dimensional framework. The resulting w values for the test statistic, change-point estimates, and p-values are reported in Table 3.

Model	U_n	W_n	$\hat{\tau}$	p-value
NL	74.92	12.43	47	0.0000
SL	36.13	7.81	45	0.0008
LL	6.60	1.79	46	0.2829
NL, SL	92.09	13.29	44	0.0000
NL, LL	89.07	13.02	47	0.0000
SL, LL	38.83	7.41	45	0.0012
NL, SL, LL	104.12	14.02	46	0.0000

Table 3. Test statistics U_n and W_n, change-point estimate $\hat{\tau}$, and corresponding p-values for the univariate, bivariate, and trivariate models for detecting change in mean only in the global annual precipitation series in Figure 1.

The results in the above table can be used in the sense that if the trivariate model was insignificant (p-value greater than 0.05), then one ignores the outcome from the bivariate and univariate models and concludes that no change is present in any of the three variables. On the other hand, when the trivariate model shows significance, then one looks for evidence from the bivariate models. In-significance in any of the bivariate models will be indicative of the fact that the pair involved did not show significant change and thus one does not pursue the univariate evidence any further for that particular pair. If all the three pairs show significance, then one concludes about significance based on the evidence from the three univariate models. Results in Table 3 show strong evidence for the trivariate and all three of the bivariate models. Hence, identification of changes in precipitation is based on the evidence from the three univariate models. Specifically, we find from the univariate models in Table 3 significant changes in mean precipitations of NL and SL latitudes, but not in the mean precipitation of the LL latitude.

Before proceeding with estimation of the unknown change-point, we need to perform model diagnostics that include validation of the three assumptions underlying the model, namely, multivariate normality, independence, and constancy of the variance-covariance matrix for the data from the three zones. For purposes of cross-validating the assumptions at the multivariate level, we utilize the residuals computed from the best fit trivariate model. The assumption of multivariate normality was tested through skewness and kurtosis tests (Mardia (1970)) , as well as a t-test (Henze and Zirkler (1990)), and the corresponding p-values were 0.063, 0.699, and 0.326, respectively. Clearly, there is no indication of violation in multivariate Gaussianity. As for independence over time, we first tested each of the three residual series from the trivariate model for significance through ACF and PACF plots up to first twenty lags and found none to be significant. We then computed the cross-correlations for each pair of residuals and found that they were also not significant. Thus, based on residuals from the change-point model, there was no indication that the assumption of independence over time was in violation. Finally, the constancy of variance-covariance structure was verified by implementing the likelihood ratio change detection statistic that tests for a change in variance-covariance matrix of the residuals. In this case, the change detection statistic U_n^* and its asymptotic distribution, which is based on W_n^* are analogous to (3.1) and (3.2) and the exact forms of both U_n^* and W_n^* may be derived from Csörgö and Horváth (1997). Here we report in Table 4 the p-values for test of change in variance based on residuals of each of the univariate, bivariate and trivariate models.

Model	U_n^*	W_n^*	$\hat{\tau}$	p-value
NL	5.20	1.29	78	0.4240
SL	1.73	-0.39	13	0.7930
LL	8.45	2.39	71	0.1681
NL, SL	5.93	0.44	78	0.7671
NL, LL	14.85	2.92	71	0.1019
SL, LL	8.94	1.42	71	0.3850
NL, SL, LL	17.25	3.63	71	0.0518

Table 4. Test statistics U_n^* and W_n^*, change-point estimate $\hat{\tau}$, and corresponding p-values for the univariate, bivariate, and trivariate models for detecting change in variance-covariance only in model residuals.

Clearly, all p-values in the above table do not provide any significant evidence against the constancy of the variance-covariance matrix. Thus, all three assumptions of Gaussianity, independence, and constancy of variance- covariance have been satisfactorily addressed in the above cross-validation process. We shall now move forward with estimating the unknown change-point through the model that best describes the change-point in the data.

First, we notice that a significant change has been detected in data from NL and SL zones whereas no significant change was found in data from the LL zone. Thus, for purposes of estimating the unknown change-point, we consider the bivariate model that involves NL and SL zones as our final change-point model. Based on this model, it is clear from Table 3 that the unknown change-point is estimated as $\hat{\tau} = 44$. No further changes in the mean were found in the data from NL and SL zones for the years 1901-1944 (p-value $= 0.6776$), or in the data for the years 1945-2000 (p-value $= 0.0832$). Thus on the basis of the change-point analysis, the mean vectors before and after the change-point for NL and SL data are estimated as $(1.5009, 1.9116)$, and $(1.5598, 2.0100)$, respectively. Also, the common mean for the LL data is estimated as 4.0493. The variance-covariance matrix for the data from the three zones is estimated by combining the residuals of the bivariate change-point model for NL and SL, and the no change model for the LL data. The estimated variance-covariance matrix and the corresponding correlation matrix for the residuals from NL, LL, and SL are presented below:

$$\begin{bmatrix} 0.0008 & -0.0010 & -0.0001 \\ -0.0010 & 0.0149 & 0.0006 \\ -0.0001 & 0.0006 & 0.0057 \end{bmatrix}, \quad \begin{bmatrix} 1.0000 & -0.2895 & -0.0468 \\ -0.2895 & 1.0000 & 0.0651 \\ -0.0468 & 0.0651 & 1.0000 \end{bmatrix}.$$

We shall now compute asymptotic distribution of the change-point mle via (3.3) and also the conditional solution (cmle) of (Cobb (1978)) via (3.4) for the bivariate change-point model involving NL and SL data. The computed distributions are presented in Table 5.

k	mle	cmle
-10		
-9		
-8		
-7	0.0001	
-6	0.0003	
-5	0.0008	
-4	0.0021	
-3	0.0064	
-2	0.0214	
-1	0.0887	0.0004
0	0.7631	0.1890
1	0.0887	0.1754
2	0.0214	0.1607
3	0.0064	0.1877
4	0.0021	0.0717
5	0.0008	0.0398
6	0.0003	0.0106
7	0.0001	0.0623
8		0.0605
9		0.0359
10		0.0032
11		0.0020
12		0.0008
13		
14		

Table 5. Probability distribution of ξ_∞ in the case of mle, and the corresponding distribution for Cobb's cmle when the normalized amount of change is $\eta = 2.48$, computed from the bivariate change-point model for NL and SL latitudes.

The 96% confidence interval based on the distribution of the mle is obtained as (1943–1946) and the same 96% confidence interval based on Cobb's cmle distribution is given by (1944–1953). Evidently, the interval based on the mle is far tighter than the one based on the cmle, and we consider (1943–1946) as our final 96% confidence interval for the change-point in the precipitation data from NL and SL zones.

Based on the above change-point analysis and the corresponding parameter estimates before and after the change-point, we find that there has been a significant increase in the mean precipitation in the NL and SL zones and that this increase has occurred subsequent to the time period (1943–1946). However, mean precipitation for the LL zone remained steady throughout the past century. The original data series together with their respective fitted models are depicted in Figure 2.

In Section 7, we shall discuss possible reasons why a change has been identified in precipitation data from NL and SL zones, whereas no change has been identified from the LL zone, and some related issues in the next section.

6. Model selection

In this section, we wish to pursue an important question regarding model fitting and model selection for the data on global precipitation. In Section 5, we pursued change-point analysis for the data on global precipitation, mainly through the mle. The change-point model was determined to be appropriate based on cross-validation of the assumptions of normality and independence over time through residual analysis. Even so, a basic question is, whether the change-point model is the right choice for the data from the view point of model selection. This question is relevant because as time series, the data on global precipitation from NL and SL latitudes exhibit serial correlations as seen through the corresponding ACF and PACF plots (not presented). Under the circumstances, one may ask why not fit a time series model, say an ARMA model that is consistent with the behavior of ACF and PACF models. The next question would then be whether the change-point models proposed in Section 5 for NL and SL data sets are any better than the corresponding best fit ARMA models? We settle this question by computing the well known Akaike Information Criterion (AIC) number for both the change-point and the best fit ARMA models. As is well known, the AIC can be applied to compare non-nested models and thus, it serves as an excellent tool for choosing an appropriate parsimonious model.

Since the data from LL (tropical) latitude did not exhibit any serial correlations and even a change-point was not detected (for this data), in this section, we restrict our attention to the data from NL and SL latitudes only. For purposes of this section, let the data from northern latitude be denoted by $Y_{N,1}, Y_{N,2}, ..., Y_{N,10}$, and correspondingly the data from southern latitude be $Y_{S,1}, Y_{S,2}, ..., Y_{S,10}$. The ACF and PACF plots suggest that an autoregressive model of order 3 ($AR\,(3)$) would be most appropriate for both data sets. Omitting all computational details, the best fit models for NL and SL latitudes were found to be:

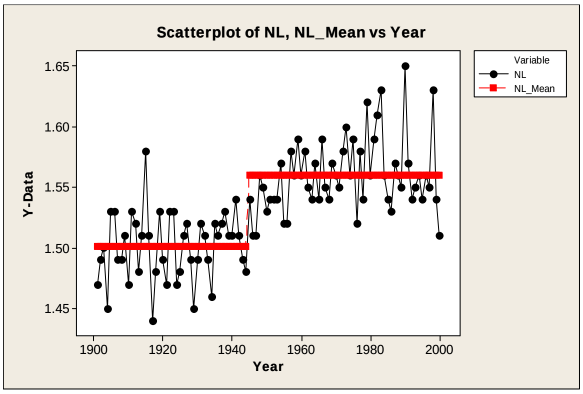

Figure 1(a) — Northern Latitude (NL)

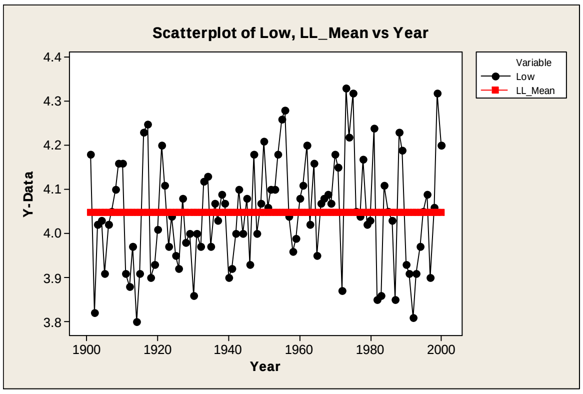

Figure 1(b) — Low Latitude (LL)

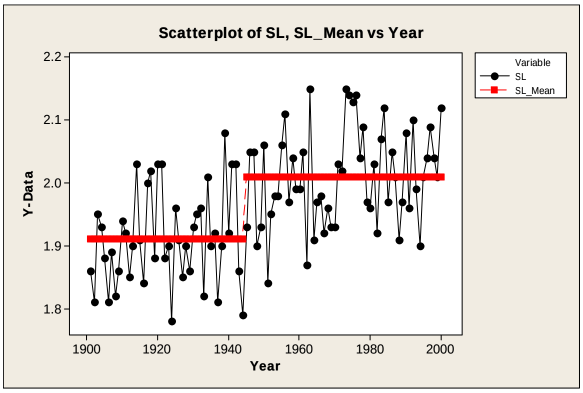

Figure 1(c) — Southern Latitude (SL)

Figure 2. Annual mean precipitation data for NL, LL, and SL zones of the globe for the years 1901-2000 together with the corresponding fitted changepoint models. (*a*) Precipitation for NL zone. (*b*) Precipitation for LL zone. (*c*) Precipitation for SL zone.

$$Y_{N,i} = 0.4173Y_{N,i-1} - 0.0256Y_{N,i-2} + 0.4095Y_{N,i-3} + \varepsilon_{N,i}, \;\; i = 4, 5, ...100, \tag{6.1}$$

$$Y_{S,i} = 0.2416Y_{S,i-1} - 0.1422Y_{S,i-2} + 0.2948Y_{S,i-3} + \varepsilon_{S,i}, \;\; i = 4, 5, ...100. \tag{6.2}$$

Once again, omitting all computational details, we provide below the AIC numbers directly for the two models above as well as for the corresponding change-point models with the change-point estimated as $\hat{\tau} = 44$ for both northern and southern latitudes.

AIC number for AR(3) model for NL data: $-681 \cdot 387$
AIC number for change-point model for NL data: $-692 \cdot 837$
AIC number for AR(3) model for SL data: $-502 \cdot 051$
AIC number for change-point model for NL data: $-512 \cdot 575$

Since one chooses the model with lowest AIC number, the above AIC numbers clearly suggest that in both NL and SL cases, the change-point model is the preferred model over the AR(3) model. Thus, the above analysis removes any ambiguity regarding the relevance of change-point model over the autoregressive model as a suitable way of modeling and analyzing the data on precipitation from and NL and SL latitudes.

7. Discussion and concluding remarks

This article implements a formal statistical analysis of the data on mean annual precipitation for NL, LL, and SL zones by adapting the statistical methodology of change-point analysis. The methodology enabled us to identify time instances subsequent to which significant and persistent changes have occurred in the precipitation over time. In particular, the analysis identified a significant increase in the global annual mean precipitation in the NL and SL zones, whereas, it remained steady throughout the past century at the LL latitudes. The 96% confidence interval for the change-point based on the distribution of the mle is found to be (1943–1946). The objective of this section is to review climatological literature that may be supportive of what we find through our analysis. We shall also provide some possible reasons that may have led to significant change in the precipitation in the NL and SL zones, but no change in the LL zone.

An excellent overview on climate trends and climate change in global as well as regional data may be found in (Dore (2005)). As per the precipitation record and the dataset used in this paper, this article reports that mean annual land precipitation has increased within the past century, even though the increase was not uniform in space and time. Some areas became drier, while others became wetter over various periods of time. The Southern and Northern hemispheres showed increased precipitation, as do the high latitudes. A precipitation increase has also been noted by several other authors. For example, Lau and Wu (2007) report that there has been an increase in precipitation since the 1950s in high latitudes, covering both the NL and SL bands. New *et al.* (2001) state that global land precipitation has gone up over the past century, with a peak noted in 1950-1960s. They observed that a large part of the increase in precipitation in the Northern Hemisphere may be due to an improved measuring gauge introduced in the 1950s. Also, New *et al.* (2001) state that in the 40-60°S (part of SL band), below average precipitation was observed before the 1930s, which was then followed by above average precipitation in the period 1930-1960s. In their regional survey on precipitation from New Zealand (part of the SL band), Plummer, Salinger, Nicholls, Suppiah, Hennessy, Leighton, Trewin, Page, and Lough (1999) assert that serious droughts were observed in the period from 1920-1951 than the following 30 year period. Below normal precipitation in the 1920-1940s was reported by Qian and Zhu (2001).

Overall, we learn from Dore (2005) that there has been an increase in mean precipitation around 1950 in the Canadian prairies, northern Europe, and west China. In the Southern Hemisphere, Australia experienced a decrease in precipitation in some regions and increased precipitation in others, while Argentina has been experiencing a positive trend in precipitation throughout the past century. On the other hand, parts of Africa have suffered severe droughts in the latter half of the century in contrast to the increases in precipitation noticed elsewhere. In the tropics and sub-tropics no overall trend has been detected thus far.

There may be other reasons why we did not detect a change in the mean tropical precipitation. First, a large part of the tropics is water and our dataset is based on land precipitation data only. Second, according to New *et al.* (2001), the LL band covers areas where precipitation has increased and areas where precipitation has decreased. For example, there are some indications that in the 20-40° S band the precipitation has increased, but in the 10° N/S to 20° N/S bands precipitation has decreased. It may be that those changes cancelled each other out and thus no overall change is detected in the aggregated dataset.

Acknowledgments

The authors thank an anonymous referee for insightful comments that have led to substantial improvements in the article. The research work of Jandhyala and Fotopoulos has been supported by National Science Foundation under grant number DMS 0806133. Jandhyala thanks CR Rao AIMSCS, Hyderabad for its partial support in carrying out this research work.

References

Berliner L, Levine R, Shea D (2000). “Bayesian Climate Change Assessment.” *Journal of Climate*, **13**, 3805–3820.

Briggs W (2008). “On the Changes in the Number and Intensity of North Atlantic Tropical Cyclones.” *Journal of Climate*, **13**, 1387–1402.

Buda S, Tong J, Guoyu R, Zhenghong C (2007). “Trends of Extreme Precipitation over the

Yangtze River Basin of China in 1960-2004." *Advances in Climate Change Research*, **3**, 45–50.

Chen J, Gupta A (2000). *Parametric Statistical Change Point Analysis.* Birkhäuser, New York, USA.

Choi K, Kim D, Byun H (2009). "The Regime Shift in the Early 1980s of Spring Precipitation in Korea." *International Journal of Climatology*, **29**. Doi: 10.1002, joc. 1927.

Cobb G (1978). "The Problem of the Nile: Conditional Solution to a Cchange-point Problem." *Biometrika*, **65**, 243–251.

Csörgö M, Horváth L (1997). *Limit Theorems in Change-point Analysis.* Wiley, New York, USA.

DeGaetano A (2006). "Attributes of Several Methods for Detecting Discontinuities in Temperature Series: Prospects for a Hybrid Homogenization Procedure." *Journal of Climate*, **9**, 1646–1660.

Dore M (2005). "Climate Change and Changes in Global Precipitation Patterns: What Do We Know?" *Environment International*, **31**, 1167–1181.

Easterling D, Peterson T (1995). "A New Method for Detecting Undocumented Discontinuous in Climatological Time Series." *International Journal of Climatology*, **15**, 369–377.

Evans D (2001). "The U.S. Climate Change Research Initiative (CCRI): Survey of Research Strategies to Reduce Scientific Uncertainties." *Technical report*, Department of Commerce.

Evans J (2009). "21st Century Climate Change in the Middle East." *Climatic Change*, **92**, 417–432.

Fearly R, Sweeny J (2005). "Detection of a Possible Change Point in Atmospheric Variability in the North Atlantic and its Effect on Scandinavian Glacier Mass Balance." *International Journal of Climatology*, **25**, 1819–1833.

Fotopoulos S, Jandhyala V, Khapalova E (2010). "Exact Asymptotic Distribution of the Change-point MLE for Change in the Mean of Gaussian Sequences." *Annals of Applied Statistics*, **4**, 1081–1104.

Giannini A, Biasutti M, Held I, Sobel A (2008). "A Global Perspective on African Climate." *Climatic Change*, **90**, 359–383.

Gonzalez-Rouco J, Jimenez J, Quesada V, Valero F (2001). "Quality Control and Homogeneity of Precipitation Data in the Southwest of Europe." *Journal of Climate*, **14**, 964–978.

Goubanova K, Li L (2007). "Extremes in Temperature and Precipitation around the Mediterranean Basin in an Ensemble of Future Climate Scenario Simulations." *Global and Planetary Change*, **57**, 27–42.

Gutman N, Hosking J, Wallis J (1993). "Regional Precipitation Quantile Values for the Continental United States Computed from L-moments." *Journal of Climate*, **6**, 2326–2340.

Hegerl G, Karl T, Allen M, Bindoff N, Gillett N, Karoly D, Zhang X, Zwiers F (2006). "Climate Change Detection and Attribution: Beyond Mean Temperature Signals." *Journal of Climate*, **19**, 5058–5077.

Hegerl G, North G (1997). "Comparison of Statistically Optimal Approaches to Detecting Anthropogenic Climate Change." *Journal of Climate*, **10**, 1125–1133.

Henze N, Zirkler B (1990). "A Class of Invariant and Consistent Tests for Multivariate Normality." *Communications in Statistics: Theory and Methodology*, **19**, 3595–3617.

Hinkley D (1972). "Time Ordered Classification." *Biometrika*, **59**, 509–523.

Ho C, Lee J, Ahn M, Lee H (2003). "A Sudden Change in Summer Rainfall Characteristics in Korea during the Late 1970s." *International Journal of Climatology*, **23**, 117–128.

Hubbard K, Flores-Mendoza F (1995). "Relating United States Crop Land Use to Natural Resources and Climate Change." *Journal of Climate*, **8**, 329–335.

Hulme M (1995). "Estimating Global Changes in Precipitation." *Weather*, **50**, 34–42.

Hurrel J (1995). "Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation." *Science*, **269**, 676–679.

Jandhyala V, Liu P, Fotopoulos S (2009). "River Stream Flows in the Northern Que'bec Labrador Region: A Multivariate Change Point Analysis via Maximum Likelihood." *Water Resources Research*, **45**. W02408, doi:10.1029, 2007WR006499.

Jaruskova (1996). "Change-point Detection in Meteorological Measurement." *Monthly Weather Review*, **124**, 1535–1543.

Karl T, Knight R, Easterling D, Quayle R (1996). "Indices of Climate Change for the United States." *Bulletin of the American Meteorological Society*, **77**(2), 279–292.

Karl T, Quayle R, Groisman P (1993). "Detecting Climate Variations and Change: New Challenges for Observing and Data Management Systems." *Journal of Climate*, **6**, 1482–1494.

Kioutsioukis I, Melas D, Zeferos C (2009). "Statistical Assessment of Changes in Climate Extremes over Greece (1955-2002)." *International Journal of Climatology*, **29**. Doi: 10.1002, joc. 2030.

Koning A, Franses P (2005). "Are Precipitation Levels Getting Higher? Statistical Evidence for the Netherlands." *Journal of Climate*, **18**, 4701–4714.

Kripalani R, Kulkarni A, Sabade S, Khandekar M (2003). "Indian Monsoon Variability in a Global Warming Scenario." *Natural Hazards*, **29**, 189–206.

Lau K, Wu H (2007). "Detecting Trends in Tropical Rainfall Characteristics, 1979-2003." *International Journal of Climatology*, **27**, 979–988.

Lund R, Reeves J (2002). "Detection of Undocumented Changepoints: A Revision of the Two-phase Regression Model." *Journal of Climate*, **15**, 2547–2554.

Mardia K (1970). "Measures of Multivariate Skewness and Kurtosis with Applications." *Biometrika*, **57**, 519–530.

Menne M, Williams J (2005). "Detection of Undocumented Changepoints Using Multiple Test Statistics and Composite Reference Series." *Journal of Climate*, **18**, 4271–4286.

Mitchell T, Carter T, Jones P, Hulme M, New M (2004). "A Comprehensive Set of High-resolution Grids of Monthly Climate for Europe and the Globe: The Observed Record (1901âĂŞ2000) and 16 Scenarios (2001âĂŞ2100)." *Tyndall Working Paper*, **55**, 1–30.

Nastos P, Zaferos C (2007). "On Extreme Daily Precipitation Totals at Athens, Greece." *Advances in Geoscience*, **337**, 1013–1022.

New M, Todd M, Hulme M, Jones P (2001). "Precipitation Measurements and Trends in the Twentieth Century." *International Journal of Climatology*, **21**, 1899–1922.

Perarnaud V, Seguin B, Malezieux E, Deque M, Loustau D (2005). "Agrometeorological Research and Applications Needed to Prepare Agriculture and Forestry to 21st Century Climate Change." *Climatic Change*, **70**, 319–340.

Plummer N, Salinger M, Nicholls N, Suppiah R, Hennessy K, Leighton R, Trewin B, Page C, Lough J (1999). "Changes in Climate Extremes over the Australian Region and New Zealand over the Twentieth Century." *Climatic Change*, **42**, 183–202.

Qian W, Zhu Y (2001). "Climate Change in China from 1880 to 1998 and its Impact on the Environmental Condition." *Climatic Change*, **50**, 419–444.

Reeves J, Chen J, Wang X, Lund R, Lu Q (2007). "A Review and Comparison of Changepoint Detection Techniques for Climate Data." *Journal of Applied Meteorology and Climatology*, **46**, 900–915.

Rodionov S (2004). "A Sequential Algorithm for Testing Climate Shifts." *Geophysical Research Letters*, **31**. L09204, doi:10.1029, 2004GL019448.

Sherwood S (2007). "Simultaneous Detection of Climate Change and Observing Biases in a Network with Incomplete Sampling." *Journal of Climate*, **20**, 4047–4062.

Smith J, Lazo J (2001). "A Summary of Cclimate Change Iimpact Assessments for the U.S. Country Studies Program." *Climatic Change*, **50**, 1–29.

Stevens A (2011). "Factors Affecting Global Climate." *Education Knowledge*, **2**(1), 5.

Tao F, Yokozawa M, Hayashi Y, Lin E (2005). "A Perspective on Water Resources in China: Interactions between Climate Change and Soil Degradation." *Climatic Change*, **68**, 169–197.

Wang X (2003). "Comments on "Detection of Undocumented Changepoints: A Revision of the Ywo-phase Regression Model "." *Journal of Climate*, **16**, 3383–3385.

Wang X (2008a). "Accounting for Autocorrelation in Detecting Mean Shifts in Climate Data Series Using the Penalized Maximal t or F Test." *Journal of Applied Meteorology and Climatology*, **47**, 2423–2444.

Wang X (2008b). "Penalized Maximal F test for Detecting Undocumented Mean Shift Without Trend Change." *Journal of Atmospheric and Oceanic Technology*, **25**, 368–384.

Wang X, Wen Q, Wu Y (2007). "Penalized Maximal t Test for Detecting Undocumented Mean Change in Climate Data Series." *Journal of Applied Meteorology and Climatology*, **46**, 916–931.

Zhao X, Chu P (2006). "Bayesian Multiple Change-point Analysis of Hurricane Activity in the Eastern North Pacific: a Markov Chain Monte Carlo Approach." *Journal of Climate*, **19**, 916–931.

Zhou Y, Xu K, Sud Y, Betts A (2011). "Recent Trends of the Tropical Hydrological Vycle Inferred from Global Precipitation Climatology Project and International Satellite Cloud Climatology Project data." *Journal of Geophysical Research*, **116**. D09101, doi:10.1029, 2010JD015197.

Affiliation:

Elena A Khapalova
University of Michigan-Flint
Flint, Michigan 48502-1950
E-mail: elenak@umflint.edu

Venkata K. Jandhyala (*Corresponding author*)
Department of Mathematics
Washington State University
Pullman, WA 99164-3113
Telephone: 509/335-8450
Fax: 509/335-1188
E-mail: jandhyala@wsu.edu

Stergios B. Fotopoulos
Washington State University
Pullman, WA 99164-4736
E-mail: fotopo@wsu.edu

Journal of Environmental Statistics http://www.jenvstat.org
Volume 4, Issue 3 Submitted: 2012-10-02
February 2013 Accepted: 2013-01-23