Special Issue Paper

Environmetrics

Received: 1 October 2011, Revised: 30 January 2012, Accepted: 17 April 2012, Published online in Wiley Online Library: 19 June 2012

(wileyonlinelibrary.com) DOI: 10.1002/env.2146

Changepoint detection in daily precipitation data†

Colin Gallagher[a], Robert Lund[a]* and Michael Robbins[b]

This paper introduces a method to identify an undocumented changepoint time in a daily precipitation series. A two-state Markov chain is used to induce dependence in the precipitation amounts; our dynamics allow for seasonality in the daily observations, a structure inherent to many nonequatorial region series. No current precipitation changepoint techniques exist that consider day-to-day dependencies, the zero support set aspect (the fact that most measurements are zero), and the periodic dynamics of the problem. The test statistic is constructed by applying cumulative sum methods to a strategically devised set of one-step-ahead prediction residuals. The methods are robust to distributional assumptions, requiring only seasonal mean and transition probability estimators. Simulations are presented that demonstrate the efficacy of the methods; application to two daily precipitation series is made. Copyright © 2012 John Wiley & Sons, Ltd.

Keywords: at most one changepoint; Brownian bridge; changepoints; CUSUM; Markov chain; precipitation

1. INTRODUCTION

Mean shift changepoints are found in many climatic series and can confound or degrade statistical inferences. United States temperature stations average about six gauge and location changes per century (Mitchell, 1953). For accurate trend assessments, it is imperative that all nonnatural discontinuities in the series be removed (or otherwise accounted for) before trends are computed. These discontinuities include effects induced by changes in station gauges or location. Climatologists are by now well aware of the changepoint issue (see Buishand, 1982; Vincent, 1998; Peterson *et al.*, 1998; Ducré-Robitaille *et al.*, 2003; Degaetano, 2006; Reeves *et al.*, 2007). An extreme example where antipodal trend conclusions are drawn when changepoints are ignored/accounted for is discussed in Lu and Lund (2007). Precipitation is, perhaps, more important than temperature for agricultural and recreational purposes. Although many authors have studied precipitation changepoints (see Buishand, 1982, Alexandersson, 1986) by looking at annual amounts, a method that accounts for some of the important features (autocorrelation and seasonality) in daily data has yet to be developed. In fact, the current state of the art for detecting daily precipitation changepoints (Wang *et al.*, 2010) simply neglects all dry days and applies Box–Cox transformations to the wet day measurements to induce Gaussianity (where standard normal homogeneity tests and their variants can be applied). No currently used precipitation changepoint test accounts for the autocorrelation found in daily precipitation data. When dependence in positively correlated data is ignored, it is easy to infer that too many changepoints exist. Lund *et al.* (2007) presents a comprehensive study of this issue. Although annually aggregating a daily precipitation series helps remove autocorrelation and frequently produces stationary (nonperiodic) data, this procedure drastically shortens the series length (and hence detection power). Moreover, one cannot infer a changepoint time to a resolution of less than a year.

To handle stochastic dependence and seasonality in the daily data, we employ a model used by Woolhiser and Pegram (1979) (see also Gabriel and Neumann, 1962; Hann *et al.*, 1976; Roldán and Woolhiser 1982; Rajagopalan *et al.*, 1996). This model can be thought of as a two-state Markov chain with periodic dynamics. The chain serves to induce dependence in the day-to-day precipitation amounts. Specifically, the chain has two states that we view as dry or wet. Conditional that a day is wet, the amount of precipitation that falls is modeled as a positive random variable with a seasonally dependent mean. Elaborating, the rainfall amounts for a wet day occurring on 5 July (for example) of any year are distributionally equivalent, but this distribution is not necessarily the same as that for 8 December wet days. Allowing for such a seasonal cycle is important for stations with a definitive dry (or monsoon) season. Precipitation over much of the United States, in fact, displays seasonal characteristics (see pages 66 and 67 of Chapman and Sherman, 1978). Whereas changepoint detection for autocorrelated and periodic temperature series was considered in Lund *et al.* (2007), issues here are slightly more delicate because of the "zero support set feature" (i.e, non-Gaussianity) of the precipitation amounts.

The test developed here is an at most one changepoint (AMOC) test. Authors have frequently adapted AMOC methods into multiple changepoint segmenters with a simple segmentation algorithm, but better alternatives exist (Davis *et al.*, 2006; Lu *et al.*, 2010). Phrased another way, our crux here lies with AMOC tests. Reference station aspects, which would likely require additional sophistication, are also not considered here. A reference station for a daily precipitation series is a similar record from a geographically nearby station that can be

* *Correspondence to: Robert Lund, Department of Mathematical Sciences, Clemson University, Clemson, SC 29634-0975, U.S.A. E-mail: lund@clemson.edu*

a *Department of Mathematical Sciences, Clemson University, Clemson, SC 29634-0975, U.S.A.*

b *Department of Statistics, University of Missouri, Columbia, MO, 65211-6100, U.S.A.*

† *This article is published in Environmetrics as a special issue on Advances in Statistical Methods for Climate Analysis, edited by Peter Guttorp, University of Washington, Norwegian Computing Center, Stephan R. Sain, National Center for Atmospheric Research, Christopher K. Wikle, University of Missouri.*

Copyright © 2012 John Wiley & Sons, Ltd.

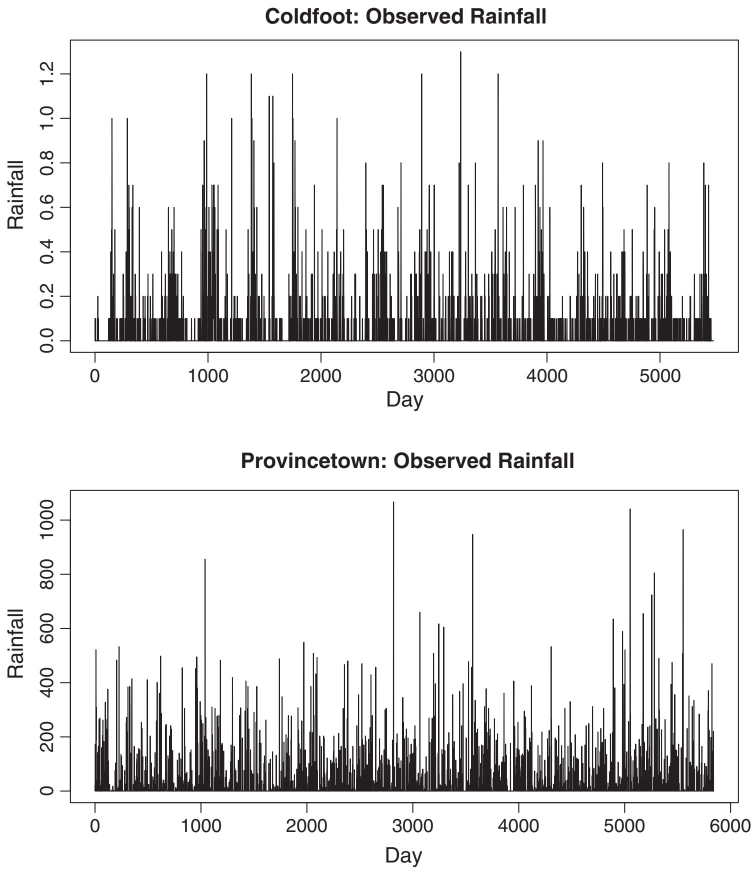

Figure 1. Coldfoot, Alaska (top) and Provincetown, Massachusetts (bottom) precipitation data. The Coldfoot units are in inches, and the Provincetown units are in tenths of millimeters

used to help illuminate any changes in the series under study. Even with these concessions, the reader will quickly gain appreciation for the complexities that arise in our setting.

Figure 1 displays two daily precipitation series that are analyzed in Section 5. The series were recorded at Coldfoot, Alaska and Provincetown, Massachusetts. Coldfoot, an aptly named Arctic hamlet, claims North America's lowest (by one degree) observed temperature of $-82°$ F. Although this record is not recognized because of gauge siting deficiencies, Coldfoot's climate is very seasonal with little precipitation falling during the extremely cold winter months. Increasing precipitation is generally conjectured to accompany a warming Arctic. Sixteen years of data from 1995–2010 are analyzed. The Provincetown series is also 16 years long and was recorded from 1953–1968. Provincetown, which resides on the tip of Cape Cod, does not have a pronounced seasonal precipitation cycle. However, there is a meta-data record for the Provincetown series. In particular, the station is known to have moved about three miles (from 42.083° latitude and $-70.216°$ longitude to 42.050° degrees latitude and $-70.183°$ degrees longitude) on 1 May 1958. In Section 5, we will see if this changepoint induced changes in the series. To obtain an exact period of 365 days, observations occurring on leap days (i.e., February 29) were deleted from both series.

The rest of this paper proceeds as follows. The next section introduces our daily precipitation model. Section 3 then devises a cumulative sum (CUSUM)-based statistic for testing whether or not a change occurs in the precipitation dynamics. The change can be either in the wet day precipitation amounts or in the background wet/dry Markov chain. Section 4 contains a simulation study that demonstrates the efficacy of the techniques. Section 5 details application of the methods to the Coldfoot and Provincetown series. The paper concludes with some comments, including some avenues for future work.

2. MODEL

We begin by describing how dependence is modeled in the daily precipitations. The model is taken from Rajagopalan *et al.* (1996). Days are divided into two types, wet and dry. A "dry" day has no recorded precipitation, and a "wet" day has a nonzero precipitation. Let $\{S_t\}$ be a process that can only be zero or one for each t. The value $S_t = 1$ indicates a wet day and $S_t = 0$ a dry day. We posit that $\{S_t\}$ is a time-homogeneous periodic Markov chain on the two states (dry and wet) with season ν one-step-ahead transition probability matrix $\mathcal{P}(\nu)$ defined by

$$\mathcal{P}(\nu) = \begin{bmatrix} \alpha(\nu) & 1-\alpha(\nu) \\ 1-\beta(\nu) & \beta(\nu) \end{bmatrix}$$

 Copyright © 2012 John Wiley & Sons, Ltd.

Here, ν represents the season being considered, and our notation takes $\nu \in \{1, \ldots, T\}$, where $T = 365$ is the period. Specifically, $\nu = 1$ corresponds to 1 January, and $\nu = 365$ corresponds to 31 December. The quantity $\alpha(\nu) \in (0, 1)$ is the probability that a season ν day (call this today) is dry given that yesterday was dry. Analogously, $\beta(\nu) \in (0, 1)$ is the probability that today, say a season ν day, is wet given yesterday was wet.

To impart seasonal features into the aforementioned dynamics, we assume that $\alpha(t)$ and $\beta(t)$ obey the first-order Fourier expansions

$$\alpha(t) = A_\alpha + B_\alpha \cos\left(\frac{2\pi(t - \tau_\alpha)}{T}\right), \quad \beta(t) = A_\beta + B_\beta \cos\left(\frac{2\pi(t - \tau_\beta)}{T}\right) \tag{2.1}$$

Higher-order Fourier or wavelet expansions could be pursued; however, day-to-day changes in precipitation are not climatologically abrupt, and the simple first-order expansions earlier usually work well.

Precipitation amounts are modeled as a positive independent sequence. Specifically, when day t is rainy, the amount of precipitation is a positive random variable X_t with mean $\mu(t)$; X_t and X_s are independent when $t \neq s$ and independent of $\{S_t\}$. Hence, all dependence in the model is induced by the two-state Markov chain. Two-state chains converge at a geometric rate to their stationary distribution; as such, our model does not have any long-memory features. The variance $\mathrm{Var}(X_t) = \sigma^2(t)$ may be periodic. Our methodology is designed to enable flexibility in the modeling of wet day rainfall amounts; hence, discussion of specific models for $\mu(t)$ and $\sigma^2(t)$ will be given later. The methods in the following do not require the marginal distribution of the wet day precipitation amounts to be specified—only the first two moments are needed.

Let Y_t be the observed precipitation on day t. The precipitation process $\{Y_t\}$ obeys

$$Y_t = S_t X_t \tag{2.2}$$

where $\{S_t\}$ follows the aforementioned Markov chain structure. This structure will be used to derive a set of one-step-ahead prediction residuals in the next section.

3. METHODS

Our methods will look for shifts in the first-order parameters of the precipitation process, which could stem from changes (changes here refer to changes beyond the natural seasonal cycle) in wet day precipitation means $\mu(\nu)$ or changes in $\alpha(\nu)$ or $\beta(\nu)$. As a likelihood ratio test requires specific distributional assumptions and would thereby be difficult to develop with seasonal dynamics and autocorrelation, a changepoint test based on CUSUM techniques is devised.

3.1. CUSUM methods

Consider locating a mean shift in an otherwise homogeneous sequence $\{\varepsilon_t\}$. If a mean shift in truth occurred at day (time) k, then a statistic comparing $\bar{\varepsilon}_k = k^{-1}\sum_{t=1}^{k}\varepsilon_t$ to $\bar{\varepsilon}_k^* = (n-k)^{-1}\sum_{t=k+1}^{n}\varepsilon_t$ should reveal differences. Here, n denotes the number of days on record. The CUSUM statistic defined by

$$\mathrm{CUSUM}_\varepsilon(k) = \frac{1}{\sqrt{n}}\left(\sum_{t=1}^{k}\varepsilon_t - \frac{k}{n}\sum_{t=1}^{n}\varepsilon_t\right) = \frac{k}{n}\left(1 - \frac{k}{n}\right)\left[\sqrt{n}\left(\bar{\varepsilon}_k - \bar{\varepsilon}_k^*\right)\right]$$

is a scaled difference of these two sample means weighting for the number of observations before and after time k. The magnitude of the maximum absolute discrepancy

$$M_\varepsilon = \max_{1 \leq k \leq n} |\mathrm{CUSUM}_\varepsilon(k)|$$

provides a mean shift changepoint test statistic, and $\mathrm{argmax}_k|\mathrm{CUSUM}_\varepsilon(k)|$ estimates the changepoint time.

An exact distribution of M_ε under a null hypothesis of no changepoint is usually intractable; however, asymptotic guidance is available. For instance, under fairly general conditions (Csörgő and Horváth, 1997)

$$M_\varepsilon \xrightarrow{\mathcal{D}} \tau \sup_{0<t<1} |B(t)| \tag{3.1}$$

where τ is a positive constant to be discussed in the following text, $\{B(t)\}_{t=0}^{1}$ denotes a Brownian bridge process, and $\xrightarrow{\mathcal{D}}$ indicate convergence in distribution. If $\{\varepsilon_t\}$ is a short-memory stationary process with lag h autocovariance $\gamma(h) = \mathrm{Cov}(\varepsilon_t, \varepsilon_{t+h})$, then

$$\tau^2 = \sum_{h=-\infty}^{\infty} \gamma(h) \tag{3.2}$$

(Berkes *et al.*, 2009). If $\{\varepsilon_t\}$ is an independent and identically distributed (IID) sequence, then $\tau^2 = \mathrm{Var}(\varepsilon_1)$. Short memory here means that $\sum_{h=0}^{\infty}|\gamma(h)| < \infty$.

 Copyright © 2012 John Wiley & Sons, Ltd.

A changepoint test based solely on (3.1) and (3.2) simply applies CUSUM techniques to the raw data $\{Y_t\}$ without worry of periodicities or zero support set issues. Such a test is investigated in Section 4. This said, periodic dynamics are important in our setup, and the convergence stated in (3.1) is slow in the presence of substantial positive autocorrelation (Robbins *et al.*, 2011). If instead the one-step-ahead prediction residuals described later are CUSUMed, temporal dependence will be accounted for and the test statistic will be sensitive to changes in $\mu(t)$ and/or $\alpha(t)$ and/or $\beta(t)$.

3.2. One-step-ahead prediction residuals

Several authors have considered changepoint testing in time series frameworks and have suggested applying CUSUM methods to estimated one-step-ahead prediction residuals (Bai, 1993, 1994; Yu, 2007; Lund *et al.*, 2007; Robbins *et al.*, 2011). Given a series $\{Y_t\}$ with finite second moments, the standardized one-step-ahead prediction residual at time t is

$$I_t = \frac{Y_t - \tilde{Y}_t}{\sqrt{\text{Var}(Y_t - \tilde{Y}_t)}} \tag{3.3}$$

Here, $\tilde{Y}_t = E(Y_t | Y_{t-1}, Y_{t-2}, \ldots, Y_1)$ is the conditional expectation prediction of Y_t from the history $Y_{t-1}, Y_{t-2}, \ldots, Y_1$, and $\text{Var}(Y_t - \tilde{Y}_t)$ is its unconditional error variance. One does not need to include the normalizing denominator $\text{Var}(Y_t - \tilde{Y}_t)^{1/2}$ in many stationary process settings; however, our dynamics are periodic, and the factor here serves to scale all prediction residuals to a common (unit-variance) basis.

For the aforementioned model, the best predictor of $Y_{iT+\nu}$ given the past, by independence of $\{X_t\}$ and $\{S_t\}$ and the Markov property of $\{S_t\}$, is

$$\begin{aligned}\tilde{Y}_{iT+\nu} &= E[X_{iT+\nu}]E[S_{iT+\nu}|Y_{iT+\nu-1}] \\ &= \mu(\nu)\left\{1_{[Y_{iT+\nu-1}>0]}\beta(\nu) + 1_{[Y_{iT+\nu-1}=0]}(1-\alpha(\nu))\right\}\end{aligned} \tag{3.4}$$

where the index $i = 0, 1, 2, \ldots$ denotes the year. Equation (3.4) holds when $iT + \nu \geqslant 2$. The startup condition is $\tilde{Y}_1 = \mu(1)P[S_1 = 1]$. To obtain $P[S_1 = 1]$, an initial distribution for S_1 needs to be specified. For this, we assume that the chain starts in its periodic stationary state (a periodic stationary state exists for a two-state chain whenever $\alpha(\nu)$ and/or $\beta(\nu)$ are not zero or one). Our notation will use $\vec{\pi}(\nu) = (\pi_d(\nu), \pi_w(\nu))'$ to denote the stationary season ν dry/wet probabilities, that is, $P[S_{iT+\nu} = 0] = \pi_d(\nu)$ and $P[S_{iT+\nu} = 1] = \pi_w(\nu)$ are constant for all i.

Identifying $\vec{\pi}(\nu)$ from the seasonal transition probabilities is not overly difficult. As the periodic stationary distribution will also arise in the variance formulas in the following text, we cannot bypass this computation. Periodic stationarity implies that $\vec{\pi}(\nu-1)\mathcal{P}(\nu) = \vec{\pi}(\nu)$ for each $\nu \in \{1, 2, \ldots, T\}$, which shows how to recursively compute $\vec{\pi}(\nu)$ for $\nu \geqslant 1$ from $\{\mathcal{P}(\nu)\}_{\nu=1}^{T}$ once $\vec{\pi}(0) = \vec{\pi}(T)$ is known. To find $\vec{\pi}(0)$, one simply solves the stationary equation $\vec{\pi}(0)\left(\prod_{\nu=1}^{T}\mathcal{P}(\nu)\right) = \vec{\pi}(0)$ subject to the restraint $\pi_d(0) + \pi_w(0) = 1$. To calculate $\vec{\pi}(0)$, one may use

$$\vec{\pi}(0)' = \mathbf{1}'\left(\boldsymbol{I} - \prod_{\nu=1}^{T}\mathcal{P}(\nu) + \boldsymbol{J}\right)^{-1}$$

where $\mathbf{1}$ is a length-2 vector of 1 s, $\boldsymbol{I}$ is the 2×2 identity matrix, and $\boldsymbol{J}$ is a 2×2 matrix of 1 s. For more on stationary distributions of periodic Markov chains, see Fralix *et al.* (2012) and the references therein.

Another computation is needed to identify $\text{Var}(\tilde{Y}_t - Y_t)$. The projection theorem gives

$$\text{Var}\left(\tilde{Y}_{iT+\nu} - Y_{iT+\nu}\right) = E\left[Y_{iT+\nu}^2\right] - E\left[\tilde{Y}_{iT+\nu}^2\right]$$

and it is easy to see that $E\left[Y_{iT+\nu}^2\right] = \pi_w(\nu)\left[\sigma^2(\nu) + \mu(\nu)^2\right]$. Using (3.4) gives

$$\tilde{Y}_{iT+\nu}^2 = \mu(\nu)^2\left\{1_{[Y_{iT+\nu-1}>0]}\beta(\nu)^2 + 1_{[Y_{iT+\nu-1}=0]}(1-\alpha(\nu))^2\right\}$$

Taking an expectation in the aforementioned equation now produces

$$\begin{aligned}\text{Var}(Y_{iT+\nu} - \tilde{Y}_{iT+\nu}) &= \pi_w(\nu)\left(\sigma^2(\nu) + \mu(\nu)^2\right) \\ &\quad - \mu(\nu)^2\left[\pi_w(\nu-1)\beta(\nu)^2 + \pi_d(\nu-1)(1-\alpha(\nu))^2\right]\end{aligned} \tag{3.5}$$

To test for a shift at an unknown time, we simply CUSUM the standardized one-step-ahead prediction residuals $\{I_t\}$ defined in (3.3). If the model dynamics do not change, then $\{I_t\}$ is a sequence of martingale differences with unit variance. This is all that is needed for CUSUM limit theory to hold (Billingsley, 1999)—independence is not needed. Our CUSUM statistic for the I_ts at time k is

$$\text{CUSUM}(k) = \frac{1}{\sqrt{n}}\left(\sum_{t=1}^{k} I_t - \frac{k}{n}\sum_{t=1}^{n} I_t\right) \tag{3.6}$$

 Copyright © 2012 John Wiley & Sons, Ltd.

and the previous work shows that

$$\max_{1 \leqslant k \leqslant n} |\text{CUSUM}(k)| \xrightarrow{\mathcal{D}} \sup_{0 \leqslant t \leqslant 1} |B(t)| \tag{3.7}$$

There is, however, one complication. The parameters governing the model are not known but need to be estimated. Before quantifying the limiting behavior of a test statistic on the basis of estimated parameter values, we describe model parameter estimation.

3.3. Parameter estimation

Consider estimating the parameters in (2.1) involving $\alpha(\nu)$. It is convenient to reexpress (2.1) in the equivalent form

$$\alpha(t) = A_\alpha + C_\alpha \cos\left(\frac{2\pi t}{T}\right) + D_\alpha \sin\left(\frac{2\pi t}{T}\right)$$
$$\beta(t) = A_\beta + C_\beta \cos\left(\frac{2\pi t}{T}\right) + D_\beta \sin\left(\frac{2\pi t}{T}\right) \tag{3.8}$$

One expansion is converted to the other via $B_\alpha^2 = C_\alpha^2 + D_\alpha^2$ and $\tau_\alpha = T \tan^{-1}(D_\alpha/C_\alpha)/2\pi$.

The harmonic coefficients for $\alpha(\nu)$ in (2.1) are estimated as follows. First, obtain an "unsmoothed estimate" of $\alpha(\nu)$ for a fixed season ν as the empirical ratio

$$\hat{\alpha}(\nu) = \frac{\{\#i : S_{iT+\nu-1} = 0 \cap S_{iT+\nu} = 0\}}{\{\#i : S_{iT+\nu-1} = 0\}} \tag{3.9}$$

(take the ratio as zero should the denominator be zero). Hence, $\hat{\alpha}(\nu)$ is simply the empirical proportion of times that a dry day in season $\nu - 1$ is followed by another dry day. Now, smooth these estimators with a harmonic fit, minimizing the weighted sum of squares

$$\sum_{\nu=1}^{T} N_\alpha(\nu) \left[\hat{\alpha}(\nu) - A_\alpha - C_\alpha \cos(2\pi\nu/T) - D_\alpha \sin(2\pi\nu/T)\right]^2 \tag{3.10}$$

in A_α, C_α, and D_α to obtain null hypothesis estimators of these parameters. Here, the weight $N_\alpha(\nu)$ is the denominator in (3.9). These weights place more emphasis on seasons with larger counts. A similar tactic working with seasonal counts of wet days followed by wet days will estimate $\beta(\nu)$.

Estimation of $\mu(t)$ and $\sigma^2(t)$ depends on how the analyst wishes to model the wet day periodicities. For instance, if $\mu(t)$ is first-order Fourier periodic and $\sigma^2(t)$ is constant in t, one can fit the linear model

$$X_t = A_\mu + C_\mu \cos\left(\frac{2\pi t}{T}\right) + D_\mu \sin\left(\frac{2\pi t}{T}\right) + \varepsilon_t \tag{3.11}$$

where $\{\varepsilon_t\}$ is IID zero mean noise to the wet day precipitation amounts. Elaborating, we time-order our wet day precipitations as $X_1, \ldots, X_p$, set $P = \{1 \leqslant t \leqslant n : S_t = 1\}$, and find A_μ, C_μ, and D_μ that minimize

$$\sum_{t \in P} \left[X_t - A_\mu - C_\mu \cos(2\pi t/T) - D_\mu \sin(2\pi t/T)\right]^2 \tag{3.12}$$

It is not clear how one can model the error sequence $\{\varepsilon_t\}$ in a pragmatic fashion while ensuring that each X_t is positive. Although this is not necessarily troublesome for simple applications of our method, in many circumstances (including our simulation study in the following text), it is easier to work with the logarithms of the precipitation amounts. Applying (3.11) and (3.12) to $\log(X_t)$ provides estimators of $\mu_\ell(t) := E[\log(X_t)]$. Further, if X_t is log-normally distributed with parameters $\mu_\ell(t)$ and σ_ℓ^2, then

$$\mu(t) = e^{\mu_\ell(t) + \sigma_\ell^{2/2}}, \quad \sigma^2(t) = \left(e^{\sigma_\ell^2} - 1\right) e^{2\mu_\ell(t) + \sigma_\ell^2} \tag{3.13}$$

Estimated values of $\mu(t)$ and $\sigma^2(t)$ are obtained by substituting estimators of $\mu_\ell(t)$ and σ_ℓ^2 into (3.13). Such estimators are $\sqrt{n}$-consistent (this follows from $\sqrt{n}$-consistency and asymptotic normality of $\hat{\mu}_\ell(t)$ and $\hat{\sigma}_\ell^2$ and a delta-method).

3.4. A Limit theorem

The time t *estimated* standardized one-step-ahead prediction residual is

$$\hat{I}_t = \frac{Y_t - \hat{Y}_t}{\sqrt{\widehat{\text{Var}}(Y_t - \tilde{Y}_t)}}$$

 Copyright © 2012 John Wiley & Sons, Ltd.

where $\hat{Y}_t$ and $\widehat{\text{Var}}(Y_t - \tilde{Y}_t)$ are versions of (3.4) and (3.5), respectively, calculated using estimated parameter values. Our estimated CUSUM statistic at index k is

$$\widehat{\text{CUSUM}}(k) = \frac{1}{\sqrt{n}}\left(\sum_{t=1}^{k}\hat{I}_t - \frac{k}{n}\sum_{t=1}^{n}\hat{I}_t\right) \tag{3.14}$$

We make a technical comment. In this setting, unlike similar settings in Bai (1993), it does not hold that $\max_{1\leqslant k\leqslant n}\left|\sum_{t=1}^{k} I_t - \sum_{t=1}^{k}\hat{I}_t\right| = o_p(\sqrt{n})$. However, we will be able to show that the difference between the CUSUM expressions in (3.6) and (3.14) is asymptotically negligible, as stated in the following lemma.

Lemma 3.1. *Under H_0, it holds that*

$$\max_{1\leqslant k\leqslant n}\left|\text{CUSUM}(k) - \widehat{\text{CUSUM}}(k)\right| = \mathcal{O}_p(n^{-1/2})$$

where $\text{CUSUM}(k)$ *and* $\widehat{\text{CUSUM}}(k)$ *are defined in (3.6) and (3.14).*

Although the proof of the lemma is technical, an outline of its justification is given in the Appendix. The result motivates the use of the changepoint statistic

$$M = \max_{1\leqslant k\leqslant n}|\widehat{\text{CUSUM}}(k)| \tag{3.15}$$

The asymptotic behavior of M is stated formally in the following result.

Theorem 3.1. *Suppose that M is as in (3.15) and that all parameter estimates are $\sqrt{n}$-consistent and that H_0 holds. Then,*

$$M \xrightarrow{\mathcal{D}} \sup_{0\leqslant t\leqslant 1}|B(t)| \tag{3.16}$$

where $\{B(t)\}_{t=0}^{1}$ is a Brownian bridge.

Equation (3.16) is derived by combining (3.7) with the result of Lemma 3.1. Percentiles for the limiting distribution in (3.16) are given in Robbins *et al.* (2011) among others.

4. SIMULATIONS

This section compares the performance of the residual-based CUSUM test of Theorem 3.1 to a simpler CUSUM procedure via simulation. As a competitor to the residual test, the convergence in (3.1) is applied to the raw precipitation amounts, replete with all zeros. This requires that periodic aspects of the problem be neglected. In the time-homogeneous case, $\{Y_t\}$ is stationary with lag-h autocovariance

$$\text{Cov}(Y_t, Y_{t+h}) = \mu^2\pi_w\left[p_{w,w}^{(h)} - \pi_w\right] + \sigma^2\pi_w 1_{[h=0]}, \quad h \geqslant 0$$

where π_w denotes the wet day stationary probability and $p_{w,w}^{(h)} = P[S_{t+h} = 1|S_t = 1]$. More detailed computations give $\pi_w = (1-\beta)/(2-\alpha-\beta)$, $\pi_d = 1-\pi_w$, $\text{Cov}(Y_t, Y_{t+h}) = \mu^2\pi_d\pi_w(\alpha+\beta-1)^{|h|}$ for $|h| \geqslant 1$, and $\text{Var}(Y_t) = \mu^2\pi_w\pi_d + \sigma^2\pi_w$. The aforementioned computations use $\alpha(\nu) \equiv \alpha$, $\beta(\nu) \equiv \beta$, $\mu(\nu) \equiv \mu$, and $\sigma^2(\nu) \equiv \sigma^2$. The quantity $\rho := (\alpha+\beta-1)$ is an important parameter since it controls the autocovariance decay rate.

From (3.1) and (3.2), it holds that

$$\max_{1\leqslant k\leqslant n}\left|\frac{1}{\tau\sqrt{n}}\left(\sum_{t=1}^{k}Y_t - \frac{k}{n}\sum_{t=1}^{n}Y_t\right)\right| \xrightarrow{\mathcal{D}} \sup_{0<t<1}|B(t)| \tag{4.1}$$

where

$$\tau^2 = \sum_{h=-\infty}^{\infty}\text{Cov}(Y_t, Y_{t+h}) = \mu^2\pi_w\pi_d\left(\frac{\alpha+\beta}{2-\alpha-\beta}\right) + \pi_w\sigma^2 \tag{4.2}$$

The convergence in (4.1) holds when τ^2 is calculated using $\sqrt{n}$-consistent parameter estimators. In the following, we will compare the performance of the residual-based test to a test that uses the left-hand side of (4.1) as a test statistic. Such a test statistic has the same asymptotic distribution as M.

 Copyright © 2012 John Wiley & Sons, Ltd.

4.1. Type I errors

This section uses simulation to study the type I error α of the various tests. The tests we consider are listed as follows.

- $\widehat{CU}$(Ind)—The CUSUM test of (4.1) assuming IID data, that is, $\tau^2 = \text{Var}(Y_1)$. Further, $\text{Var}(Y_1)$ is estimated with the sample variance.
- $CU(Y)$—The CUSUM test of (4.1), where τ^2 as defined in (4.2) is used and assumed known (exact parameters are used).
- $\widehat{CU}(Y)$—The CUSUM test of (4.1), where τ^2 as defined in (4.2) is used, assumed unknown, and estimated under the null hypothesis of no changepoints.
- CU(R)—The residual CUSUM test of Theorem 3.1, where all parameters are known.
- $\widehat{CU}$(R)—The residual CUSUM test of Theorem 3.1, where parameters are unknown and are estimated under the null hypothesis.

We begin by considering the effects of autocorrelation on the type I error rates. To isolate on correlation aspects, no periodic effects are included in these simulations. Figure 2 shows simulated type I errors for various values of $\rho \in (-1, 1)$. Each type I error was aggregated from 50,000 independently generated series. The chain parameters are taken as $\alpha = \beta = (\rho + 1)/2$, and rainfall amounts were generated from a log-normal distribution with parameters $\mu_\ell = -1$ and $\sigma_\ell^2 = 0.5$. The number of years of data was $N = 10$, that is, $n = 365N$.

The results show that the residual test performs well for all magnitudes of correlation. Tests that fail to incorporate autocorrelation have little utility in settings where ρ is close to unity. Because of this, we will not consider the $\widehat{CU}$(Ind) test further. CUSUMing the raw data (while incorporating correlation aspects) appears to be a viable option in this case.

To investigate convergence issues, we vary N while using the same setup used to generate Figure 2. Here, $\rho = 0.8$ is fixed, and no periodic dynamics are used in the simulations. Empirically aggregated type I errors are reported in Figure 3. The residual test does not perform particularly well when $N = 1$ or $N = 2$. This is attributed to estimation of the seasonal parameter structure (seasonal parameters were estimated even though the dynamics chosen were not seasonal). In fact, the type I errors for the residual test using known parameters (CU(R)) are close to the nominal 5% level for all N.

Next, the performance of the tests in the presence of periodic dynamics will be explored. For chain transition probabilities, we take $A_\alpha = 0.7$, $A_\beta = 0.5$, and $D_\alpha = D_\beta = 0$ in (3.8). For the C parameters in (3.8), we vary $p \in [0, 1]$ and take $C_\alpha = -0.25p$ and $C_\beta = 0.45p$.

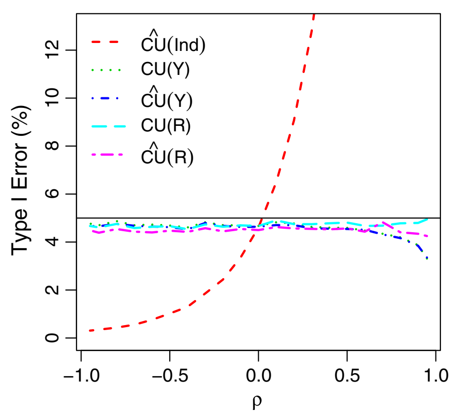

Figure 2. Simulated type I error rates for various ρ. The residual-based tests hold the correct 5% level for all ρ

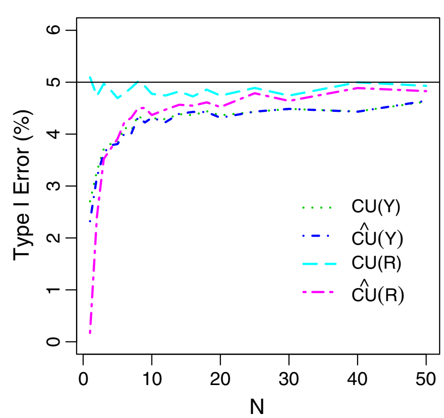

Figure 3. Simulated type I error rates for various N. The tests need a few years of data to perform well

 Copyright © 2012 John Wiley & Sons, Ltd.

A larger p entails more seasonality in the model dynamics. The precipitation amounts were generated from a log-normal distribution, where $\mu_\ell(\nu)$ and $\sigma^2(\nu)$ are defined for a log-scale version of (3.11) with $A_\mu = -0.5$, $C_\mu = p$, $D_\mu = 0$, and $\sigma_\ell^2 \equiv 0.5$. Figure 4 displays type I errors for various values of p. Again, seasonality increases with increasing p. Some of the simulated type I errors for the $CU(Y)$ and $\widehat{CU}(Y)$ tests are significantly bigger or smaller than 5%; however, the residual test's type I error stays close to the targeted 5% value. The inference is that one should not apply nonperiodic tests to series with periodic dynamics.

4.2. Detection power

This section studies the detection power of the tests. Our intent is to demonstrate that the residual test can effectively detect a single changepoint. In the simulations later, a changepoint mean shift is added to the data at time (expressed as a portion of the data) $c/n = 0.726$. For

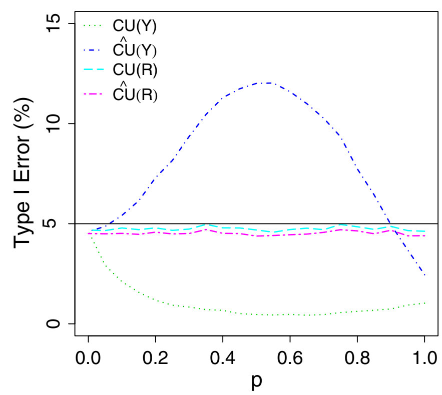

Figure 4. Simulated type I error rates for various strengths of periodicity p. The residual-based test's error rate stays close to 5%

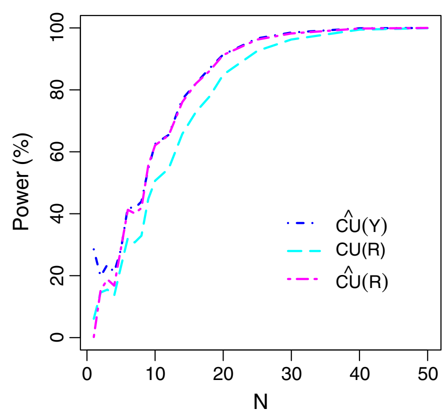

Figure 5. Simulated powers for various N with periodic data. The powers increase to unity with increasing N

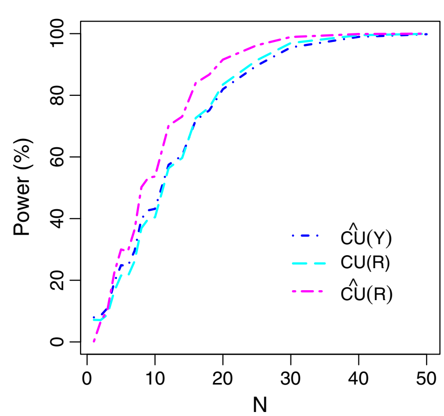

Figure 6. Simulated powers for various N with nonperiodic data. The powers increase to unity with increasing N

 Copyright © 2012 John Wiley & Sons, Ltd.

realism, the changepoint time, mean shift, and other parameters are set to those estimated for the Coldfoot series in the next section. The specific parameter values are listed in Table 2. We vary N to verify that detection powers increase to unity as the sample size increases. Estimated powers are shown in Figure 5. As exact parameters are not known in this situation, results for $CU(Y)$ cannot be presented.

Figure 5 suggests that the residual test detects the changepoint effectively when $N \geqslant 5$. Although the $\widehat{CU}(Y)$ test has larger power when $N = 1$ and $N = 2$, it also erroneously accepts changepoints too many times (as was shown in Figure 4). To assess whether the residual test is less powerful than the $\widehat{CU}(Y)$ test, we put both tests on the same footing. Elaborating, the simulations in Figure 5 are repeated after removing all periodic dynamics: $C_\alpha = D_\alpha = 0$, $C_\beta = D_\beta = 0$ and $C_\mu = D_\mu = 0$. Figure 6 displays these powers. The other parameters for the residual tests are estimated in a time-homogeneous fashion. The residual test does not appear to suffer any loss of power; in fact, for larger values of N, it appears that the residual test is more powerful than tests that directly CUSUM the data.

5. APPLICATIONS

This section examines the Coldfoot, Alaska and Provincetown, Massachusetts precipitation series of the introduction. We first look at the Coldfoot series.

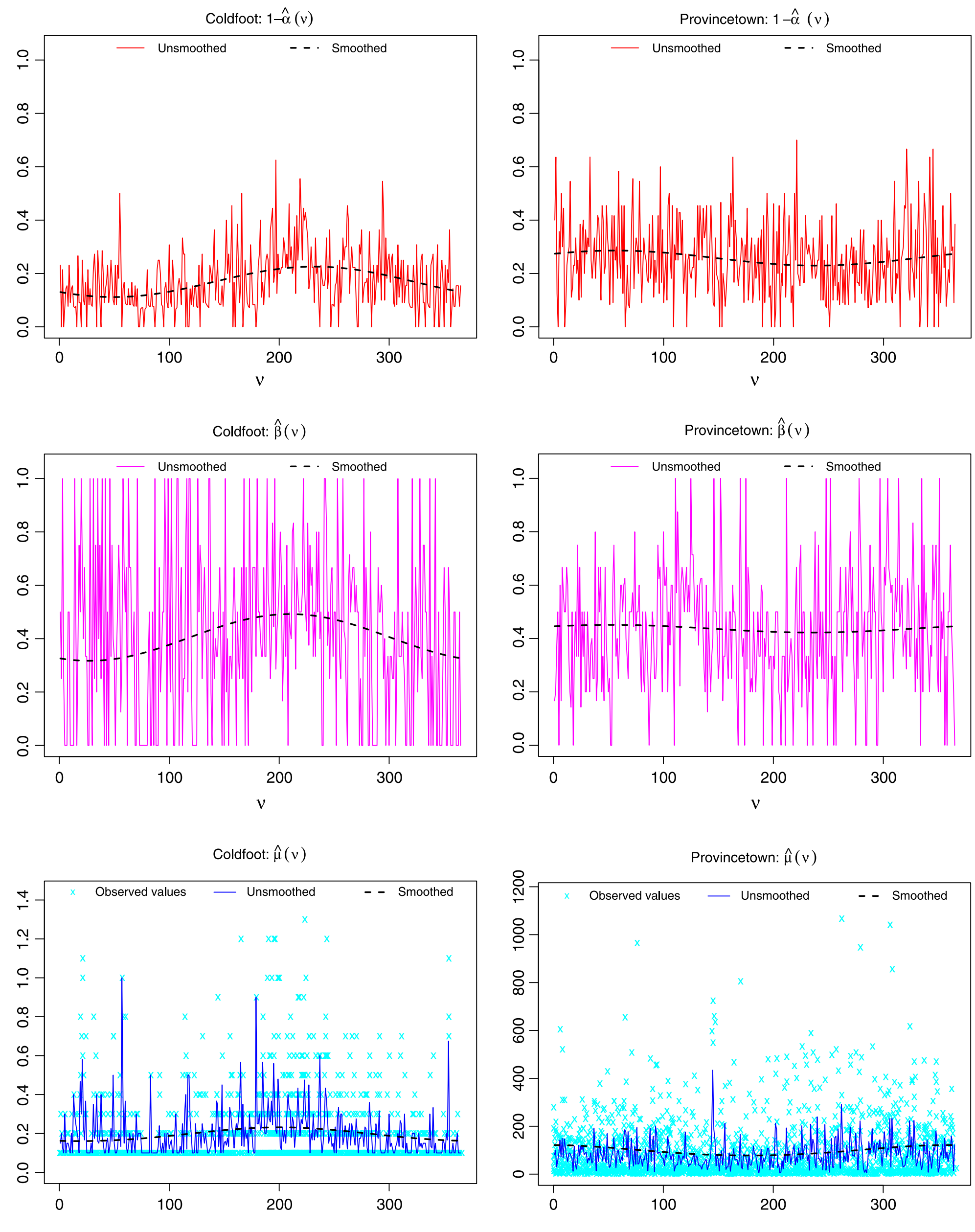

Figure 7. Plots demonstrating the periodic features (or lack thereof) of the Coldfoot (left column) and Provincetown (right column) estimated parameters. Plotted sequences include Markov chain dry/wet (top row) and wet/wet (middle row) transition probabilities and wet day mean rainfall (bottom row)

Copyright © 2012 John Wiley & Sons, Ltd.

The periodic structure of the Coldfoot data is graphically portrayed in Figure 7. The top left plot shows the empirical "unsmoothed" values of $1-\alpha(\nu)$ for $\nu \in \{1,\ldots,T\}$, as found using (3.9), along with the smoothed values found by minimizing (3.10). The middle left plot contains corresponding results for $\beta(\nu)$. The empirical version of $\beta(\nu)$ portrays more dispersion than the corresponding version of $\alpha(\nu)$ because there are fewer wet days (than dry days) on record. The bottom left plot in Figure 7 shows an unsmoothed version of $\mu(\nu)$ against the first-order Fourier mean fitted assuming the observations follow a log-normal distribution. The bottom plot also displays all positive precipitation amounts. Note that all precipitation amounts have been rounded to the nearest 0.1 in. The other two plots show estimated versions of $1-\alpha(\nu)$ and $\beta(\nu)$ (that is, the probability of rainfall on day ν after a dry day and wet day, respectively). As the structures in all plots are quite seasonal, it would be inappropriate to proceed with time-homogeneous dynamics.

The changepoint tests in Section 4 were applied to the Coldfoot data. The results are listed in Table 1. All tests locate a changepoint at time $c = 3974$ or $c = 3973$ (fall of 2006) that is significant at the 5% level. It is not clear what caused this changepoint as no meta-data is available. However, the transition has been to dryer conditions (see Figure 8). To assess the impact of outliers, all precipitation measurements greater than 0.8 in were truncated to 0.8 in, and the analysis was rerun. The p-value for the $\widehat{CU}(R)$ tests increases from 0.0132 to 0.0302.

Table 2 shows the estimated parameter values before and after the flagged changepoint time and assuming a changepoint-free scenario (i.e., H_0 holds). The changepoint alters the values of C_α, A_β, D_β, C_μ, and D_μ. Figure 8 plots observed precipitation amounts replete with all zeros against the fitted mean structure (unconditional) of the data. Although the changes in the mean at the changepoint time are not

Table 1. Results of changepoint tests for the Coldfoot and Provincetown series

	Coldfoot		Provincetown	
	$\widehat{CU}(Y)$	$\widehat{CU}$(R)	$\widehat{CU}(Y)$	$\widehat{CU}$(R)
Stat	1.4940	1.5840	0.3865	0.4403
$\hat{c}$	3974	3973	1201	3729
p-value	0.0230	0.0132	> 0.5	> 0.5

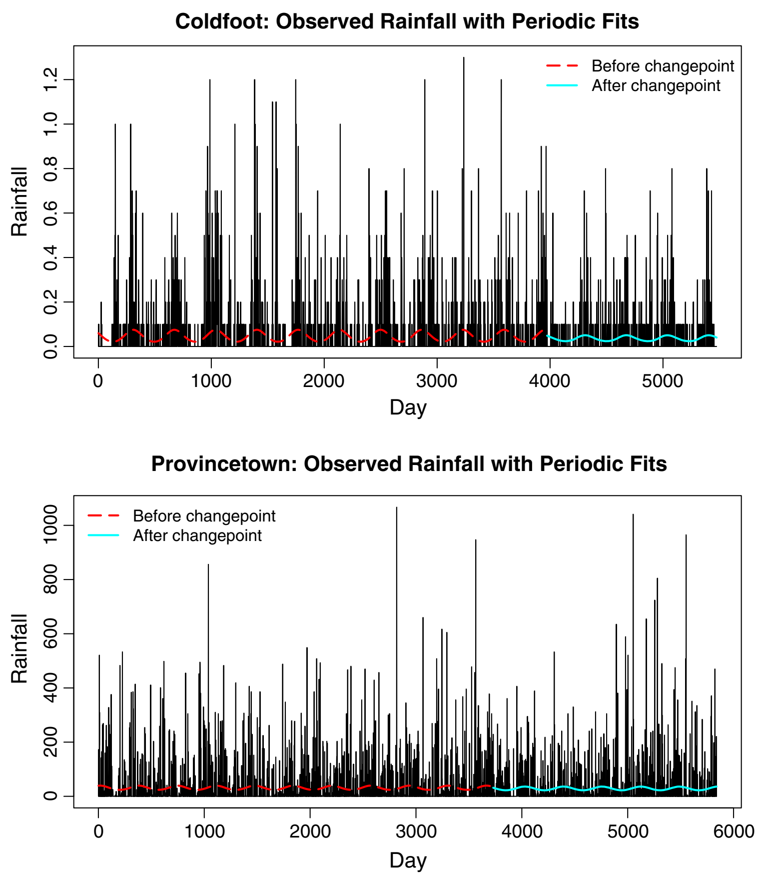

Figure 8. Observed precipitation amounts against fitted mean structure for Coldfoot (top) and Provincetown (bottom) data

 Copyright © 2012 John Wiley & Sons, Ltd.

Table 2. Estimated parameters for the Coldfoot and Provincetown series. The Coldfoot figures estimate the breakpoint at $\hat{c} = 3973$

Coldfoot										
	A_α	C_α	D_α	A_β	C_β	D_β	A_μ	C_μ	D_μ	σ_ℓ^2
Under H_0	0.8314	0.0377	0.0429	0.04046	−0.0781	−0.0397	−1.8415	−0.1737	−0.0554	0.3936
Before $\hat{c}$	0.8338	0.0491	0.0439	0.4308	−0.0696	−0.0617	−1.8084	−0.1568	−0.0742	0.4243
After $\hat{c}$	0.8238	0.0065	0.0407	0.3314	−0.0902	0.0265	−1.9306	−0.2082	−0.0003	0.2943
Provincetown										
	A_α	C_α	D_α	A_β	C_β	D_β	A_μ	C_μ	D_μ	σ_ℓ^2
Under H_0	0.7423	−0.0160	−0.0232	0.4368	0.00886	0.0113	3.5779	0.2232	−0.01748	1.9980

particularly large, it is detectable because of the large number of days on record. The changepoint has reduced the overall mean of the series; in addition, there appears to be a decrease in the amplitude of the harmonic component of the mean.

Next, the methods were applied to the Provincetown series. The measurements in this series are recorded to tenths of a millimeter. Table 1 shows that the 1 May 1958 station move does not induce a changepoint. In fact, p-values for all tests exceed 0.5. Perhaps, this is not surprising given the homogeneity of the terrain in Provincetown. The estimated model parameters in Table 2 indicate that the seasonality in the wet/dry and wet/wet one-step-ahead transition probabilities is not as pronounced as that for the Coldfoot series; this point is graphically illustrated in Figure 7. Figure 8 shows the unconditional mean of the model against the data.

Finally, a QQ plot (not shown) of the wet day precipitation amounts was made to assess log-normality of the Provincetown data (rounding issues make such an assessment difficult for the Coldfoot data). Formally, log-normality is rejected; however, as noted earlier, distributional assumptions are not needed as the methods only employ the first two seasonal moments.

6. COMMENTS

CUSUM methods can be adapted to handle changepoint detection issues in daily precipitation series. Although CUSUMing the raw data replete with zeroes is not a radically inferior approach, one must account for autocorrelations in the data to obtain reliable inferences. Also, it is not advisable to apply nonperiodic techniques to periodic data.

Modifications of our methods merit exploration. For example, the statistic

$$M^* = \max_{\ell \le \frac{k}{n} \le h} \frac{\widehat{\mathrm{CUSUM}}^2(k)}{\frac{k}{n}\left(1 - \frac{k}{n}\right)} \tag{6.1}$$

is often more powerful than M but requires that one "crops" the data boundaries to an interval $[\ell, h] \subset (0, 1)$. Specifically, one must take ℓ small and positive and h close to (but less than) unity. This statistic has close connections to likelihood ratios and has higher power than M when the changepoint is not near the center of the data record. Robbins *et al.* (2011) discussed these issues in detail. The limiting distribution of M^* under the null hypothesis of no changepoints is known to be

$$M^* \xrightarrow{\mathcal{D}} \sup_{\ell \le t \le h} \frac{B^2(t)}{t(1-t)}$$

where $\{B(t)\}_{t=0}^1$ is again a Brownian bridge process. Null hypothesis percentiles are easily obtained; however, one must select ℓ and h. Selection of these parameters is usually not crucial, and we have successfully used $\ell = 0.01$ and $h = 0.99$ in other applications.

A short simulation was conducted to compare the performances of M and M^*. We use the setup in Figure 5 with the following modifications: to increase the magnitude of the change, we use $A_\mu = -1.508$ prior to the changepoint, and we impose a more "recent" changepoint — $c/n = 0.925$, specifically. For $N = 10$ with $\ell = 0.05$ and $h = 0.95$, the power increases from 0.23658 to 0.79356.

We make one final comment. It is not clear how to handle reference station aspects. One possible method would incorporate reference station aspects via a linear regression covariate term in the one-step-ahead predictions in (3.4). Beaulieu *et al.* (2011) have recently explored similar tactics with promising results.

Acknowledgements

Robert Lund's research was supported by National Science Foundation Grant DMS 0905770. Comments from two referees and the editor helped improve this manuscript.

 Copyright © 2012 John Wiley & Sons, Ltd.

REFERENCES

Alexandersson H. 1986. A homogeneity test applied to precipitation data. *International Journal of Climatology* **6**: 661–675.
Bai J. 1993. On the partial sums of residuals in autoregressive and moving average models. *Journal of Time Series Analysis* **14**: 247–260.
Bai J. 1994. Weak convergence of the sequential empirical processes of residuals in ARMA models. *The Annals of Statistics* **4**: 2051–2061.
Beaulieu C, Chen J, Sarmiento JL. 2011. Change point analysis as a tool to detect abrupt climate variations. *Philosophical Transactions of the Royal Society, Series A* **370**(1962): 1228–1249.
Berkes I, Gombay E, Horváth L. 2009. Testing for changes in the covariance structure of linear processes. *Journal of Statistical Planning and Inference* **139**: 2044–2063.
Billingsley P. 1999. *Convergence of Probability Measures*, Second Edition. John Wiley and Sons: New York City.
Buishand TA. 1982. Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology* **58**: 11–27.
Chapman JD, Sherman JC. 1978. *Oxford Regional Economic Atlas. The United States and Canada*, Second Edition. Oxford University Press: London.
Csörgő M, Horváth L. 1997. *Limit Theorems in Change-Point Analysis*. John Wiley and Sons: Chichester.
Davis RA, Lee TCM, Rodriguez-Yam GA. 2006. Structural break estimation for nonstationary time series models. *Journal of the American Statistical Association* **101**: 223–239.
Degaetano AT. 2006. Attributes of several methods for detecting discontinuities in mean temperature series. *Journal of Climate* **19**: 838–853.
Ducré-Robitaille J-F, Vincent LA, Boulet G. 2003. Comparison of techniques for detection of discontinuities in temperatures. *International Journal of Climatology* **23**: 1087–1101.
Gabriel KR, Nuemann J. 1962. A Markov model for daily rainfall occurrence at Tel Aviv. *Quarterly Journal of the Royal Meteorological Society* **88**: 90–95.
Hann CT, Allen DM, Street JO. 1976. A Markov chain model of daily rainfall. *Water Resources Research* **12**: 443–449.
Fralix B, Livzey J, Lund RB. 2012. Renewal sequences with periodic dynamics. *Probability in the Engineering and Informational Sciences* **26**: 1–15.
Lu Q, Lund RB, Lee TCM. 2010. An MDL approach to the climate segmentation problem. *Annals of Applied Statistics* **4**: 299–319.
Lu Q, Lund RB. 2007. Simple linear regression with multiple level shifts. *Canadian Journal of Statistics* **37**: 447–458.
Lund RB, Wang XL, Reeves J, Lu Q, Gallagher C, Feng Y. 2007. Changepoint detection in periodic and autocorrelated time series. *Journal of Climate* **20**: 5178–5190.
Mitchell JM, Jr. 1953. On the causes of instrumentally observed secular temperature trends. *Journal of Applied Meteorology* **10**: 244–261.
Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Torok S, Auer I, Boehm R, Gullett D, Vincent LA, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger J, Frland EJ, Hanssen-Bauer I, Alexandersson H, Jones P, Parker D. 1998. Homogeneity adjustments of in situ atmospheric climate data: A review. *International Journal of Climatology* **18**: 1493–1517.
Rajagopalan B, Lall U, Tarboton DG. 1996. Nonhomogeneous Markov model for daily precipitation. *Journal of Hydrologic Engineering* **1**: 33–40.
Reeves J, Chen J, Wang XL, Lund RB, Lu Q. 2007. A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology* **46**: 900–915.
Robbins M, Gallagher C, Lund RB, Aue A. 2011. Mean shift testing in correlated data. *Journal of Time Series Analysis* **32**: 498–511.
Roldán J, Woolhiser DA. 1982. Stochastic daily precipitation models 1. A comparison of occurrence processes. *Water Resources Research* **18**: 1451–1459.
Vincent LA. 1998. A technique for the identification of inhomogeneities in Canadian temperature series. *Journal of Climate* **11**: 1094–1104.
Wang XL, Chen H, Wu Y, Feng Y, Pu Q. 2010. New techniques for the detection and adjustment of shifts in daily precipitation data series. *Journal of Applied Meteorology and Climatology* **49**: 2416–2436.
Woolhiser DA, Pegram GGS. 1979. Maximum likelihood estimation of Fourier coefficients to describe seasonal variations of parameters in stochastic daily precipitation models. *Journal of Applied Meteorology* **18**: 34–42.
Yu H. 2007. High moment partial sum processes of residuals in ARMA models and their applications. *Journal of Time Series Analysis* **28**: 72–91.

APPENDIX

Proof of Lemma 3.1. Let $\hat{\alpha}(t)$, $\hat{\beta}(t)$, and $\hat{\mu}(t)$ represent the estimated versions of $\alpha(t)$, $\beta(t)$, and $\mu(t)$, respectively. Recall that $\alpha(t) = \alpha(\nu)$, $\beta(t) = \beta(\nu)$, and $\mu(t) = \mu(\nu)$ whenever $t = iT + \nu$ for $\nu \in \{1, \ldots, T\}$—the same will hold for the estimators and the stationary distributions, $\{\pi_d(t), \pi_w(t)\}_{t=1}^n$. Set $\zeta_t = \text{Var}(Y_t - \tilde{Y}_t)^{-1/2}$ and use $\hat{\zeta}_t$ to represent the version of ζ_t calculated using the estimated parameter values. Define $\tilde{S}_t$ and $\hat{S}_t$ so that $\tilde{Y}_t = \mu_t \tilde{S}_t$ and $\hat{Y}_t = \hat{\mu}_t \hat{S}_t$. The $\sqrt{n}$-consistency of the parameter estimators gives

$$\max_{1 \leqslant \nu \leqslant T} ||\{\hat{\mu}(\nu), \hat{\alpha}(\nu), \hat{\beta}(\nu), \hat{\zeta}_\nu\}' - \{\mu(\nu), \alpha(\nu), \beta(\nu), \zeta_\nu\}'|| = \mathcal{O}_p\left(n^{-1/2}\right) \tag{A.1}$$

for the Euclidean norm $||\cdot||$.

Let $\epsilon_t = Y_t - \tilde{Y}_t$ be the true prediction residuals and let $\hat{\epsilon}_t = Y_t - \hat{Y}_t$ be the estimated prediction residuals. Additionally, let R_k represent the difference between the true and estimated partial sum sequences at time k. Then,

$$R_k = \sum_{t=1}^{k} (\zeta_t \epsilon_t - \hat{\zeta}_t \hat{\epsilon}_t) = \sum_{t=1}^{k} \zeta_t [\tilde{S}_t \mu(t) - \hat{S}_t \hat{\mu}(t)] + \sum_{t=1}^{k} \hat{\epsilon}_t (\zeta_t - \hat{\zeta}_t)$$

Using (3.4) and simplifying provide

$$\begin{aligned} R_k = & \sum_{t=1}^{k} \zeta_t \hat{\mu}(t) \left[\alpha(t) - \hat{\alpha}(t)\right] (S_{t-1} - \pi_w(t-1)) + \sum_{t=1}^{k} \zeta_t \hat{\mu}(t) [\beta(t) - \hat{\beta}(t)](S_{t-1} - \pi_w(t-1)) \\ & + \sum_{t=1}^{k} \zeta_t \left[\mu(t) - \hat{\mu}(t)\right] (\tilde{S}_t - \pi_w(t)) + \sum_{t=1}^{k} (\zeta_t - \hat{\zeta}_t) \hat{\epsilon}_t + \sum_{t=1}^{k} B_t \end{aligned} \tag{A.2}$$

 Copyright © 2012 John Wiley & Sons, Ltd.

where

$$B_t = \zeta_t \mu(t)[\hat{\alpha}(t) - \alpha(t)]\pi_d(t-1) + \zeta_t \mu(t)[\beta(t) - \hat{\beta}(t)]\pi_w(t-1) + \zeta_t[\mu(t) - \hat{\mu}(t)]\pi_w(t)$$

Define $A_k = R_k - \sum_{t=1}^{k} B_t$. To show that $\max_{1 \leqslant k \leqslant n} |A_k| = \mathcal{O}_p(1)$, we need to show that the first four sums in (A.2) are $\mathcal{O}_p(1)$. For the first term, it is sufficient to consider ks that are whole multiples of T. For $k = rT$,

$$\sum_{t=1}^{k} \zeta_t \hat{\mu}(t)[\alpha(t) - \hat{\alpha}(t)](S_{t-1} - \pi_w(t-1)) = \sum_{\nu=1}^{T} \zeta_\nu \hat{\mu}(\nu)[\alpha(\nu) - \hat{\alpha}(\nu)] \sum_{i=1}^{r} [S_{iT+\nu-1} - \pi_w(iT+\nu-1)]$$

For each fixed ν,

$$\max_{1 \leqslant r \leqslant N} \left| \sum_{i=1}^{r} [S_{iT+\nu-1} - \pi_w(iT+\nu-1)] \right| = \mathcal{O}_p(\sqrt{n})$$

and by (A.1), $|\zeta_\nu \hat{\mu}(\nu)[\alpha(\nu) - \hat{\alpha}(\nu)]| = \mathcal{O}_p(n^{-1/2})$. Thus, the first term is $\mathcal{O}_p(1)$. The second and third summations are handled similarly.

The arguments for the fourth term are essentially the same, except one must deal with the fact that ϵ_t is being estimated. More detailed work using the consistency of the estimators shows that

$$\max_{1 \leqslant k \leqslant n} \left| \sum_{t=1}^{k} \hat{\epsilon}_t \right| = \mathcal{O}_p(\sqrt{n})$$

and the fourth term is $\mathcal{O}_p(1)$ as well. Note that the term $\sum_{t=1}^{k} B_t$ is not $\mathcal{O}_p(1)$. In fact, $\max_{1 \leqslant k \leqslant n} |R_k| = \mathcal{O}_p(\sqrt{n})$.

Next, consider the difference between the true and estimated CUSUM statistics:

$$\sqrt{n}\left(\text{CUSUM}(k) - \widehat{\text{CUSUM}}(k)\right) = \sum_{t=1}^{k} B_t - \frac{k}{n}\sum_{t=1}^{n} B_t + A_k - \frac{k}{n} A_n \tag{A.3}$$

where CUSUM(k) and $\widehat{\text{CUSUM}}(k)$ are defined in (3.6) and (3.14). We note that the B_t are periodic in that $B_\nu = B_{iT+\nu}$ for $\nu = 1, \ldots, T$. Letting $\bar{B} = \sum_{t=1}^{T} B_t$, invoking periodicity, and using $\sum_{\nu=1}^{T}(B_\nu - \bar{B}) = 0$ give

$$\begin{aligned}
\left| \sum_{t=1}^{k} B_t - \frac{k}{n}\sum_{t=1}^{n} B_t \right| &= \left| \sum_{t=1}^{k} (B_t - \bar{B}) - \frac{k}{n}\sum_{t=1}^{n} (B_t - \bar{B}) \right| \\
&\leqslant \left| \sum_{t=1}^{k} (B_t - \bar{B}) \right| + \left| \sum_{t=1}^{n} (B_t - \bar{B}) \right| \\
&\leqslant 2T \max_{1 \leqslant t \leqslant T} |B_t - \bar{B}| \\
&\leqslant 4T \max_{1 \leqslant t \leqslant T} |B_t| = \mathcal{O}_p(n^{-1/2})
\end{aligned} \tag{A.4}$$

Using (A.3) followed by $\max_{1 \leqslant k \leqslant n} |A_k| = \mathcal{O}_p(1)$ and (A.4) proves Lemma 3.1.

 Copyright © 2012 John Wiley & Sons, Ltd.