Advances in Computer, Signals and Systems (2016) 1: 1-7
Clausius Scientific Press, Canada

Bayesian analysis of hydrological time series based on MCMC algorithm

Zhao Huiqin[1, a], Liu Jinshan[*,2,b]

[1]Hua Shang College, GuangDong University of Finance & Ecnomics, Guangzhou, China
[2]College of Mathematics and Informatics, South China Agricultural University, Guangzhou, China
[a]zhq_6285144@163.com,[b]liujs58@scau.edu.cn
[*]Corresponding Author

Keywords: Bayesian stochastic search, MCMC algorithm, Hydrological time series, Change-point regression model.

Abstract: In this paper we consider Bayesian analysis of the possible changes in hydrological time series by Markov chain Monte Carlo (MCMC) algorithm. We consider multiple change-points and various possible situations. The approach of Bayesian stochastic search selection is used for detecting and estimating the number and positions of possible change-point in a piecewise constant model. MCMC algorithm is used to estimate the posterior distributions of parameters. The result of the analysis is applied to the hydrological data sets of the major river net area of Shunde in China and the data set of Nile River. In order to further investigate the trends in each segment of the hydrological data sets, we consider the analysis of change-point regression model via MCMC algorithm.

1. Introduction

Following the published studies on climate changes, a number of hydrologists have used models, which describe certain types of changes, to represent hydrological time series. When natural surroundings changes abruptly, the hydrological time series may exhibit some trends or jumps. In this case, the statistical characteristics of the sequence may be different before and after some time points known as the change-points. Thus, change-point of hydrological time series reflects certain environmental changes. In most of the previous papers, a given type of change occurring with certainty is assumed, and focus has been put on the characterization of the change-point [1-2]. However, the problem of detecting multiple change-points is one of the most challenging problems, since both the number of the change-points and their locations are unknown. Besides the positions of the change-points, the trends of the sequence in different segments are of interest. Regression model can be used to deal with this problem.

In this paper we consider the problem of analyzing the possible changes in hydrological time series by MCMC methods. In section 2, we adopt a piecewise constant multiple change-points Bayesian procedure to analyze the real data sets of the major river net area of Shunde in China and the data set of the Nile River from [3]. In order to further investigate the trends in different segments of the time series, we consider the analysis of change-points regression model by MCMC algorithm in section 3. Finally, in section 4, we present some conclusions.

2. Bayesian change-points analysis

2.1 Piecewise constant change-points model

Let $\{y_t, t \geq 1\}$ be a real stochastic process which is constant between two change-points such that

$$y_t = m_k, \text{ for all } \tau_{k-1} + 1 \leq t \leq \tau_k, \tag{1}$$

where $\{m_k, k \geq 1\}$ is a real sequence, $\{\tau_k, k \geq 0\}$ is a set of change-points with the convention $\tau_0 = 0$, $\{\varepsilon_t, t \geq 1\}$ is assumed to be a sequence of independent Gaussian random variables such that

$$\varepsilon_t \sim N0, \sigma_k^2), \text{ for all } \tau_{k-1} + 1 \leq t \leq \tau_k. \tag{2}$$

It is convenient for detecting change-points to introduce a random vector $\gamma = (\gamma_1, ..., \gamma_{n-1})$, where

$$\gamma_t = \begin{cases} 1, & \text{if thre exists k such that } t = \tau_k, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The estimation of the change-points instants reduces to the estimation of the set $\gamma = (\gamma_1, ..., \gamma_{n-1})$ and $\{m_k, k \geq 1\}$. For a given configuration of γ, $K_\gamma = \sum_{t=1}^{n-1} \gamma_t + 1$ denotes the number of segment. Let $m = (m_1, ..., m_{K_\gamma})$ and $\sigma^2 = (\sigma_1^2, ..., \sigma_{K_\gamma}^2)$, $n_k = \tau_k - \tau_{k-1}$, then the likelihood function is given by

$$f(y \mid \gamma, m; \sigma^2) = \prod_{k=1}^{K_\gamma} (2\pi\sigma_k^2)^{-\frac{n_k}{2}} \exp\{-\frac{1}{2} \sum_{k=1}^{K_\gamma} \frac{1}{\sigma_k^2} \sum_{t=\tau_{k-1}+1}^{\tau_k} (y_t - m_k)^2\}. \tag{4}$$

In Bayesian inference, the choice of priors is important. We suppose γ is a sequence of independent and identically distributed Bernoulli random variables with parameter λ. Thus, the prior density of γ is

$$\pi(\gamma, \lambda) = \lambda^{K_\gamma - 1} (1 - \lambda)^{n - K_\gamma}. \tag{5}$$

On the other hand, for a given γ, the prior of m is chosen as independent Gaussian distributions

$$\pi\left(m \mid \gamma; \mu, V\right) = \prod_{k=1}^{K_\gamma} \left(\frac{2\pi V}{n_k}\right)^{-1/2} \exp\left\{-\frac{n_k}{2V}\left(m_k - \mu\right)^2\right\}. \tag{6}$$

Let $\theta = (\mu, \lambda, \sigma^2, V)$ be the set of hyper-parameters which are known constants. Using standard distribution theory we obtain the conditional posterior distribution of m from equation (6),

$$f\left(m \mid y, \gamma; \theta\right) = \prod_{k=1}^{K_\gamma} \left(2\pi V_k\right)^{-\frac{1}{2}} \exp\{-\left(m_k - \mu_k\right)^2 / (2V_k)\}, \tag{7}$$

where $V_k = \frac{V\sigma_k^2}{n_k(V + \sigma_k^2)}, \mu_k = \frac{V\sigma_k^2}{V + \sigma_k^2}(\frac{\overline{y}_k}{\sigma_k^2} + \frac{\mu}{V}), \overline{y}_k = \frac{1}{n_k}(\sum_{t=\tau_{k-1}+1}^{\tau_k} y_t)$. It means that conditional on (y, γ), the $m_k's$ remains independent and Gaussian.

According to Bayes theorem, the joint posterior distribution of the parameter (γ, m) can expressed as

$$f\left(\gamma, m \mid y; \theta\right) \propto \left(y \mid \gamma, m; \sigma^2\right) \pi\left(m \mid \gamma; \mu, V\right) \pi\left(\gamma; \lambda\right). \tag{8}$$

Consequently, the parameters $m_k^{'}s$ can be eliminated by integrating out m_k from equation (8). In a special case, $\sigma_k^2 = \sigma^2, k = 1,...,K_\gamma$, The conditional posterior distribution of γ can be obtained as

$$f(\gamma \mid y;\theta) = C(y;\theta)\exp\{-U(\gamma \mid y;\theta)\}. \tag{9}$$

where $C(y;\theta)$ is a normalizing constant, while $U(\gamma \mid y;\theta) = \phi S_\gamma + rK_\gamma$ is referred to as the energy function, in which $\phi = \frac{V}{2\sigma^2(V+\sigma^2)}$, $S_\gamma = \sum_{k=1}^{K_\gamma} S_k$, $r = \frac{1}{2}\ln(\frac{(V+\sigma^2)}{\sigma^2}) + \ln\frac{1-\lambda}{\lambda}$.

The unknown parameter vector γ can be estimated from the posterior distribution $f(r \mid y;\theta)$. One of the standard Bayesian estimates is the (marginal) maximum a posteriori (MAP) estimator obtained by maximizing the posterior distribution defined as $\hat{\gamma} = \text{argmax } f(\gamma \mid y;\theta) = \text{argmax } \mathrm{U}(\gamma \mid y;\theta)$. Unfortunately, a closed-form expression of the MAP estimator of γ cannot be obtained and MCMC algorithm could be used to obtain it.

2.2 MCMC methods

The main idea of this algorithm is to generate a Markov chain $\{\gamma^{(i)}, i \geq 0\}$ using Metropolis-Hastings (M-H) algorithm with the invariant distribution $f(\gamma \mid y;\theta)$. The M-H algorithm is an iterative procedure. At iteration i, we carry out the following two steps:

(1) An admissible new value $\tilde{\gamma}$ is drawn from a proposal kernel $q(\tilde{\gamma} \mid \gamma^{(i)})$, which is irreducible.

(2) $\tilde{\gamma}$ is accepted as the new state $\gamma^{(i+1)} = \tilde{\gamma}$ with the acceptance probability:

$$\alpha(\gamma^{(i)}, \tilde{\gamma}) = \min\{1, \frac{f(\tilde{\gamma} \mid y;\theta)q(\tilde{\gamma} \mid \gamma^{(i)})}{f(\gamma^{(i)} \mid y;\theta)q(\gamma^{(i)} \mid \tilde{\gamma})}\}. \tag{10}$$

Since $f(\gamma \mid y;\theta) \propto \exp\{-U(\gamma \mid y;\theta)\}$, if the proposal kernel is symmetric such that $q(\tilde{\gamma} \mid \gamma^{(i)}) = q(\gamma^{(i)} \mid \tilde{\gamma})$, then the acceptance step (2) reduces to:

$$\gamma^{(i+1)} = \begin{cases} \tilde{\gamma}, & \text{if } U(\gamma^{(i)} \mid y;\theta) - U(\tilde{\gamma} \mid y;\theta) > \eta, \\ \gamma^{(i)}, & \text{otherwise}. \end{cases} \tag{11}$$

where $\eta = \ln R$, R is drawn from the uniform distribution U(0,1). After a sufficiently long burn-in, the estimator of γ is determined by computing the time average of the Markov chain output samples.

The MAP estimator of γ is determined by using a simulated annealing (SA) algorithm, which defines a non-homogeneous Markov chain. A decreasing temperature schedule should be introduced in the SA algorithm which can modify the acceptance probability. The proposal kernels are defined as in [4]:

(a) The candidate γ is drawn independently of the current state $\gamma^{(i)}$ from an instrumental distribution defined by $q(\tilde{\gamma} \mid \gamma^{(i)}) = q(\tilde{\gamma})$. In this paper, q is chosen as a Bernoulli distribution with parameter λ.

(b) A random permutation of $\{1,...,n\}$ is uniformly drawn. According to this permutation, each component is flipped from 0 to 1 or from 1 to 0.

(c) An actual change-point is randomly selected and a neighborhood of this instant is defined. The change-point instant is moved in its neighborhood and accepted according to the acceptance probability.

The acceptance probability is given by equation (10). The schedule for lowering the temperature is defined by $T_k = 0.99T_{k-1}$, where T_0 is greater than a numerical constant depending on the energy

function $U(\gamma \mid y;\theta)$. This temperature decreases per cycle. The acceptance procedure is modified by

$$r^{(i+1)} = \begin{cases} \tilde{r}, & \text{if } U(r^{(i)} \mid y;\theta) - U(\tilde{r} \mid y;\theta) > \eta T_i, \\ r^{(i)}, & \text{otherwise,} \end{cases} \quad (12)$$

The implementation of a MCMC algorithm as described above assumes that the set of hyper-parameters θ is known. We estimate θ in a maximum likelihood framework by using the Stochastic Approximation version of the EM (SAEM) algorithm given by [5]. The algorithm describes as follows: (a) Choose an initial guess $\theta^{(0)}$ and an initial configuration of change-points instants $\gamma^{(0)}$. (b) At step j of the iteration: (1) perform one iteration of MCMC using the current value of $\theta^{(j-1)}$ to simulate $\gamma^{(j)}$ from $\gamma^{(j-1)}$. (2) Compute the maximum likelihood estimate $T(\gamma^{(j)})$ of θ by maximizing the joint distribution of (y,γ) and update $\theta^{(j)}$ by $\theta^{(j)} = \theta^{(j-1)} + \alpha_j(T(r^{(j)}) - \theta^{(j-1)})$, where $\{\alpha_j\}$ is the step size sequence which decreases to 0.

2.3 Data analysis with hydrological data

In this subsection, we apply the MCMC method to analyze fourteen data sets, in which thirteen time series are consisted of the annual maximum values in the major hydrological stations of Shunde in China. The other data set is the annual volume of the Nile River.

The algorithm used in this paper draws vectors $\gamma^{(i)}$ according to the distribution $f(\gamma \mid y;\theta)$ with the MCMC algorithm proposed by Gibbs Sampling. For each vector $\gamma^{(i)}$, the estimated number of segment is $K(\gamma^{(i)}) = 1 + \sum_{t=1}^{n-1} \gamma_t^{(i)}$. The MAP estimator of $\hat{\gamma}$ is obtained after enough iteration (150000). Finally, m is drawn with the conditional posterior distribution $p(m \mid \hat{\gamma}, y;\theta)$.

The posterior distributions of γ, i.e. the probabilities $\{P(\gamma_t = 1 \mid y;\theta), \{1 \le t \le n-1\}$, is displayed by Fig.1. The posterior distribution of K_γ, i.e. the probabilities $\{P(K_r = k \mid y;\theta), 1 \le k \le n\}$ obtained by computing the histogram of the estimated numbers $K(\gamma^{(i)})$ is displayed in Fig.2, while Fig.3 depicts the time series of the hydrological data and the estimate of the vector m. The results are summarized in Table 1 and we summarize some conclusion as follows:

(1) Only one change-point was detected for each series.

(2) MCMC method detects one change-point at 1899 for the Nile river series, which agrees with previous studied by [6].

(3) Da Zhou, Fu Zhou He, Huan Ma Yong, Ma Kou, Rong Qi, San Hong Qi, Xiao Bu, Xin Yong have the same change time of 1993. Most of them are located in a developed area. It seems that these eight downstream sections are affected by human activities more obviously.

(4) The change time of Ban Sha Wei,Gan Zhu, Nan Sha, San Shui and Wan Qing Sha are more earlier than those detected by [6] using a grey relational method.

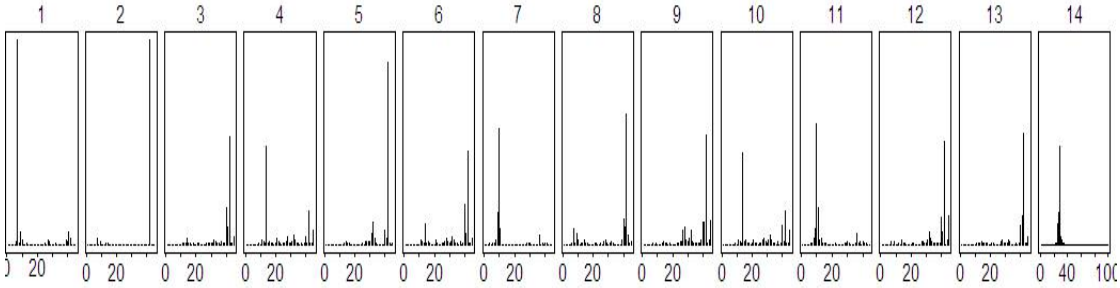

Fig. 1 The posterior distribution of γ estimated with 150 000 iterations

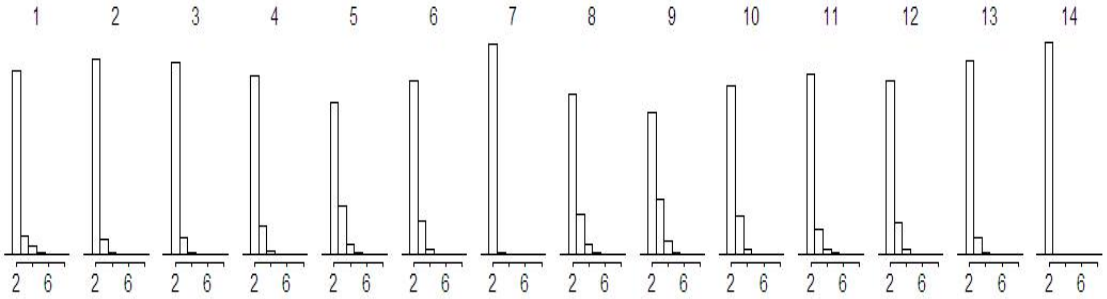

Fig. 2 The posterior distribution of number K_γ of segments estimated with 150 000 iterations

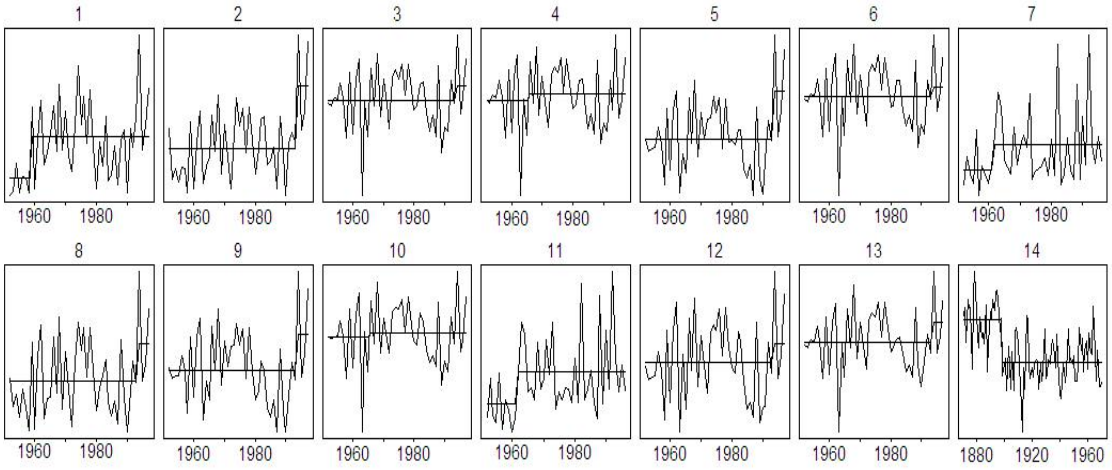

Fig. 3 The time series of the data sets of Shunde and Nile River, along with the estimate $\hat{m}$

3. Bayesian change-points regression analysis

3.1 Change-points regression model

In this section, we consider the segmented regression model with k change-points for fitting the observed data given by

$$y_t = \alpha_k + \beta_k t + \varepsilon_t \text{ , for all } \tau_{k-1} + 1 \le t \le \tau_k, 1 \le k \le K, \tag{13}$$

with the convention $\tau_0 = 0$, where $\{\varepsilon_t, t \ge 1\}$ is a sequence of random variables with normal distributions defined by equation (2). This model is also called segmented-line regression model.

Let $\gamma = (\gamma_1, ..., \gamma_{n-1})$ and K_γ be defined as in section 2. Let $\sigma^2 = (\sigma_1^2, ..., \sigma_{K_\gamma}^2)$ and $\eta = \{\eta_1, ..., \eta_{K_\gamma}\}$, where $\eta_k = (\alpha_k, \beta_k)', k = 1, 2, ..., K_\gamma$. Then, the likelihood function is given by:

$$f(y \mid r, \eta; \sigma^2) = \prod_{t=1}^{n} \left(2\pi\sigma_t^2\right)^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma_t^2}(y_t - \alpha_t - \beta_t t)^2\}. \tag{14}$$

Suppose the prior density $\pi(\gamma; \lambda)$ of γ is defined by eq. (5) and let $\eta_k, k = 1, 2, ...,$ be independently distributed as the multivariate normal distribution $N(\eta_0, \Sigma)$, where η_0 and Σ are hyper-parameters. Thus for a given configuration γ, the prior density of η is defined by:

$$\pi(\eta \mid r; \eta_0, \Sigma) = \prod_{k=1}^{K_\gamma} \pi(\eta_k \mid r; \eta_0, \Sigma) = \prod_{k=1}^{K_\gamma} (2\pi)^{-\frac{1}{2}} \mid \Sigma \mid^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\eta_k - \eta_0)'\Sigma^{-1}(\eta_k - \eta_0)\}. \tag{15}$$

For a given sequence of segment points $\{\tau_k, k \geq 1\}$, define

$$B^{(k)} = (\sigma_k^{-2} X^{(k)} X^{(k)'} + \Sigma^{-1})^{-1}, b^{(k)} = \sigma_k^{-2} X^{(k)} Y^{(k)} + \Sigma^{-1} \eta_0$$

where $k = 1,..., K_\gamma$, and

$$X^{(k)} = \begin{pmatrix} 1 & \cdots & 1 \\ \tau_{k-1}+1 & \cdots & \tau_k \end{pmatrix}, Y^{(k)} = \begin{pmatrix} Y_{\tau_{k-1}+1} \\ \vdots \\ Y_{\tau_k} \end{pmatrix}. \quad (16)$$

Let $\theta = (\sigma^2, \eta_0, \Sigma)$ denote the set of hyper-parameters. Then the conditional posterior distribution of η is given by $p(\eta \mid y, r; \theta) \propto f(y \mid r, \eta; \sigma^2)\pi(\eta \mid r; \eta_0, \Sigma)$. Using standard distribution theory we obtain the following full conditional distributions：

$$\eta_k \mid y, r; \theta \sim N(B^{(k)} b^{(k)}, B^{(k)}), k = 1,..., K_\gamma. \quad (17)$$

The marginal posterior distribution of γ is

$$f(\gamma \mid y; \theta) \propto f(y \mid \gamma, \eta; \sigma^2)\pi(\gamma; \lambda) \propto \prod_{k=1}^{K_\gamma} (\sigma_k^2)^{-\frac{n_k}{2}} \exp\{-\frac{1}{2\sigma_k^2}(Y^{(k)} - X^{(k)}\eta_k)'(Y^{(k)} - X^{(k)}\eta_k)\} \times (\frac{\lambda}{1-\lambda})^{K_\gamma}. \quad (18)$$

In the special case where $\sigma_k^2 = \sigma^2, k = 1,..., K_\gamma$, the marginal posterior distribution (18) reduces to

$$f(r \mid y; \theta) \propto (\frac{\lambda}{1-\lambda})^{K_\gamma} \exp\{-\frac{1}{2\sigma^2}\sum_{k=1}^{K_\gamma}(Y^{(k)} - X^{(k)}\eta_k)'(Y^{(k)} - X^{(k)}\eta_k)\}. \quad (19)$$

The estimation of γ could be obtained by the MCMC method discussed in section 2. For a given configuration γ, the estimation of regression parameters $\eta = (\eta_1,..., \eta_{K_\gamma})$ can be obtained from the above standard distributions.

3.2 The analysis of hydrological data

In this subsection, we use the change-points regression model via MCMC method to analyze the data sets of the annual maximum values in the data set as in section 2. In the MCMC procedure, the hyper-parameter η_0 is chosen as the least squares estimator from the regression model $y = \alpha + \beta t + \varepsilon$, while σ_k^2 is chosen as the sample variance of the observations $\{y_t, 1 \leq t \leq n\}$ for $k \geq 1$. The hyper-parametric matrix Σ is chosen as an identity matrix $I_2 = diag(1,1)$.

The sample-based parameter estimates and the posterior probabilities of the fitted models are presented in Table 1. The change-point instant τ with maximum posterior probability is the same as detected in section 2 for each data set, but the maximum posterior probability may be different from that under the piecewise constant model in section 2. The trends in each segment of the fourteen hydrological data sets can be clearly seen from the estimates of the regression coefficients in Table 1.

4. Conclusions

This paper studies the problem of change-point analysis of hydrological time series by Bayesian method. We consider various possible situations that may occur. The probabilistic model makes use

of a non-observed sequence γ, and an M-H algorithm is used for estimating the posterior distribution of γ. The advantage of this parameterization is that the dimension of the sequence γ is fixed and the hyper-parameters of the model are estimated by the SAEM algorithm.

The method of the paper is applied to analyze the hydrological data sets of the major river net area of Shunde in China and the data set of Nile River. These results agree with those obtained in the previous literatures. The MCMC method is used to estimate also the posterior distribution of the mean sequence. For further investigate the trends in each segment of the time series we use a segmented change-points regression model via MCMC approach. The change-point instant detected with maximum posterior probability is the same as that in the piecewise constant model for each data set. The trends in each segment can be clearly seen from the estimates of the segmented regression models.

Table 1 Summary

ID	rivers	K_γ	τ	Piece wise constant model		regression model		
				$P(\tilde{\gamma}\mid y,\theta)$	$\tilde{m}=(\tilde{m}_1,\tilde{m}_2)$	$P(\tilde{\gamma}\mid y,\eta)$	$\tilde{\eta}_1=(\tilde{\alpha}_1,\tilde{\beta}_1)$	$\tilde{\eta}_2=(\tilde{\alpha}_2,\tilde{\beta}_2)$
1	Ban Sha Wei	2	1958	0.802	2.0;2.4	0.5932	1.932;0.015	2.324;0.002
2	Da Zhou	2	1993	0.926	2.5;3.2	0.497	2.309;0.006	2.053;0.025
3	Fu Zhou He	2	1993	0.383	4.0;4.4	0.3146	4.043;-0.003	11.690;-0.156
4	Gan Zhu	2	1965	0.289	4.6;4.9	0.1842	4.863;-0.156	5.230;-0.011
5	HuanMa Yong	2	1993	0.508	3.1;3.2	0.5422	3.091;-0.004	5.078;-0.029
6	Ma Kou	2	1993	0.293	7.4;7.7	0.2736	7.458;-0.005	17.970;-0.214
7	Nan Sha	2	1961	0.563	1.7;1.9	0.2838	1.699;0.009	1.848;0.268
8	Rong Qi	2	1993	0.476	2.6;3.1	0.6084	2.593;0.002	8.587;-0.120
9	San Hong Qi	2	1993	0.308	3.0;3.6	0.5696	3.239;-0.011	6.047;-0.053
10	San Shui	2	1965	0.171	7.4;7.5	0.2104	7.359;-0.064	8.301;-0.021
11	Wan Qing Sha	2	1961	0.537	1.7;2.0	0.3514	1.753;-0.002	1.890;0.028
12	Xiao Bu	2	1993	0.326	4.1;4.6	0.355	4.158;-0.004	10.863;0.129
13	Xin Yong	2	1993	0.438	3.1;3.5	0.4602	3.151;-0.004	1.480;-0.153
14	Nile	2	1899	0.665	935;913	0.8236	1080.0;1.117	805.0;0.696

References

[1] Ruggieri E., Antonellis M. An exact approach to Bayesian sequential change point detection. Computational Statistics and Data Analysis Vol. 97 (2016), pp.71–86

[2] Lu K.P., Chang S.T. Detecting change-points for shifts in mean and variance using fuzzy classification maximum likelihood change-point algorithms. Journal of Computational and Applied Mathematics Vol.308 (2016), pp.447–463

[3] Cobb G. W. The problem of Nile: conditional solution to a change point problem. Biometrika Vol.65(1978), pp. 243–251

[4] Lavielle M, Lebarbier E. An application of MCMC methods for the multiple change-points problem. Siginal Processing Vol.81(2001),pp. 39–53

[5] Tourneret J. Y, Doisy M, Lavielle M. Bayesian off-line detection of multiple change-points corrupted by multiplicative noise: application to SAR image edge detection. Signal Processing Vol.83(2003), pp.1871~1887

[6] Wong H, Hu B. Q, Ip, W. C, Xia J. Change-point analysis of hydrological time series using grey relational method. J. Hydrol. Vol.324(2006), pp.323–338