ON CHANGE-POINT ANALYSIS OF MAXWELL DISTRIBUTION USING BAYESIAN TECHNIQUES

*[1]Taiwo M. Adegoke & [2]Oladapo M. Oladoja

•

[1,2]Department of Mathematics and Statistics, First Technical University, Ibadan, Nigeria.
[1]taiwo-adegoke@tech-u.edu.ng, [2]oladapo.oladoja@tech-u.edu.ng

Abstract

This research work focuses on Bayesian inference in this study to detect a change in the rate of a Maxwell distribution model with independent random variables. The paper specifically analyzes a single rate shift and demonstrates how the Bayesian framework can be used to efficiently solve this problem. To produce samples from Maxwell distribution and evaluate the datasets, simulation techniques were used, and the R programming language was used. Although the model looks to be simple, no analytical solutions are available for parameter inference, necessitating the use of approximations. The study emphasizes the Gibbs sampler's applicability for change-point analysis using a Markovian updating approach. The simulation research findings show that the predicted rate is near to the true value, confirming the consistency and stability of the Bayesian estimator.

Keywords: Gibbs sampling, Change-point, Bayes factor, Bayesian method, Conjugate prior distribution

1. Introduction

Change-point analysis (CPA) is a statistical technique for detecting and quantifying changes in data across time. CPA identifies data points when there is a significant shift in the underlying structure or behavior of the process producing the data. CPA finds applications in various domains. It is utilized in fault detection and reliability [1], insurance, econometric timeseries, and malware software detection [2]. Furthermore, it plays a role in signal detection, surveillance, security systems, meteorology, and climatology [3] . Changepoint analysis is also employed in graphical models [4], gynecology [5], communication network evolution [6, 7, 8], oceanography [9], sparse VAR models [10, 11], macrosociological processes and historical changes [12], medicine [13], and functional magnetic resonance recordings [14, 15] , among others.

CPA can be traced back to the work of [16, 17, 18], where cumulative sums (CUMSUM) approach was used to identify points of change in a sequence of normally distributed observations. Since then, several methods have been proposed for performing CPA, including Bayesian methods (see [19, 20, 21, 22, 23, 24, 25], likelihood-based methods (see [26, 27, 28, 29, 30, 31] and non-parametric methods (see [32, 33, 34, 35]). These techniques involve various statistical models and algorithms to estimate the change points accurately. The choice of method depends on the characteristics of the data and the specific objectives of the analysis.

In their study,[36] investigated the change point analysis (CPA) in the Maxwell distribution using Bayesian methods. They examined both informative and non-informative prior distributions and utilized two distinct loss functions, namely Linex LF and General Entropy LP, to detect the change point and estimate its magnitude. However, the present work introduces a novel approach by utilizing Bayes' factor techniques to detect a single change point in a series of observations following a Maxwell distribution. The method utilizes a conjugate prior distribution and employs a Monte Carlo Gibbs sampling approach to estimate the parameters involved.

2. Method

2.1. Bayesian Techniques

The definition of a change-point, as proposed by [16, 17, 18] and [26], involves a test to determine whether a sequence of independent observations, arranged in a successive order $x_1, x_2, \ldots, x_n$, are drawn from the same probability density function $F(x|\theta)$, which is characterized by the likelihood function.

$$L(x;\theta) = \prod_{i=1}^{n} f(x;\theta) \tag{1}$$

as against set of observations with a single change-point k represented as $x_1, x_2, \ldots x_k$ and $x_{k+1}, x_{k+2}, \ldots x_n$ before and after the change having two distinct probability density functions $F(x|\theta_1)$ and $F(x|\theta_2)$, where $\theta_1 \neq \theta_2$. The likelihood function for the alternate hypothesis can be expressed as

$$L(x;\theta_1, \theta_2) = \prod_{i=1}^{k} f(x;\theta_1) \prod_{i=\gamma+1}^{n} f(x;\theta_2) \tag{2}$$

In the Bayesian perspective, a joint prior distribution $p(\theta_1, \theta_2)$ is assumed for the parameters. Bayes theorem then provides the joint posterior distribution

$$p(\theta_1, \theta_2, \gamma | x, y) = \frac{p(x, y|\theta_1, \theta_2)p(\theta_1, \theta_2)}{\int \int p(x, y|\theta_1, \theta_2)p(\theta_1, \theta_2)\partial\theta_1\partial\theta_2} \tag{3}$$

The prior distribution $p(\theta_1, \theta_2)$ reflects the beliefs about the parameters before experimentation, whereas the posterior distribution $p(\theta_1, \theta_2, k|x)$ reflects the updated beliefs about the parameters after observing the sample data.

2.2. Bayes' Factor

Bayesian statisticians perceive hypothesis testing as a process of comparing models ([37], [38]) rather than focusing on whether a specific hypothesis is true, the emphasis is placed on determining which model, described under one hypothesis, is more favorable compared to another. The Bayesian approach to hypothesis testing was initially developed by [39, 40] as a fundamental component of scientific inference. A central aspect of Jeffreys' framework involved the concept of the *Bayes factor*, which represents the posterior odds of the null hypothesis when the prior probability on the null is one-half. Jeffreys employed this approach to compare predictions made by two competing scientific theories. In this methodology, statistical models are introduced to represent the likelihood of the observed data according to each theory, and Bayes' theorem is utilized to compute the posterior probability that one theory is correct.

In their study, [37] consider a dataset D, which is assumed to be generated under two hypotheses: H_1 and H_2. The probability densities $\zeta(D/H_1)$ and $\zeta(D/H_2)$ describe the data under each hypothesis, respectively. Prior probabilities, $\zeta(H_1)$ and $\zeta(H_2) = 1 - \zeta(H_1)$, are assigned to H_1 and H_2, respectively. By applying Bayes' theorem, the authors obtain the posterior probabilities $\zeta(H_1/D)$ and $\zeta(H_2/D)$ as follows:

$$\zeta(H_i/D) = \frac{\zeta(D/H_i)\zeta(H_i)}{\zeta(D/H_1)\zeta(H_1) + \zeta(D/H_2)\zeta(H_2)}, \quad (i = 1, 2) \tag{4}$$

$$\zeta(H_1/D) = \frac{\zeta(D/H_1)\zeta(H_1)}{\zeta(D/H_1)\zeta(H_1) + \zeta(D/H_2)\zeta(H_2)}$$

$$\zeta(H_2/D) = \frac{\zeta(D/H_2)\zeta(H_2)}{\zeta(D/H_1)\zeta(H_1) + \zeta(D/H_2)\zeta(H_2)}$$

In certain applications, such as testing hypotheses regarding the presence of a change-point, it is often more informative to consider the odds in favor of H_2 compared to H_1 ([41, 42]).

$$\frac{\zeta(H_2/D)}{\zeta(H_1/D)} = \frac{\zeta(D/H_2)}{\zeta(D/H_1)} \times \frac{\zeta(H_2)}{\zeta(H_1)} \tag{5}$$

and the transformation is simply multiplication of the prior odds by

$$B_{12} = \frac{\zeta(D/H_2)}{\zeta(D/H_1)} = \frac{\int_{\theta_2} \zeta(D/\theta_2)\zeta(\theta_2)\mathrm{d}\theta_2}{\int_{\theta_1} \zeta(D/\theta_1)\zeta(\theta_1)\mathrm{d}\theta_1} \tag{6}$$

which is the baye's factor. Thus,in words,
posterior odd = bayes factor $\times$ prior odds
and the bayes factor is the ratio of the posterior odd of H_1 to its prior odds, regardless of the value of the prior odds.

By analogy with the likelihood ratio obtained from Equation (6) (i.e the quantity $\log B_{12}$) is often used to summarize the evidence for H_2 compare to H_1, with the rough interpretation shown in Table 1. This contrasts with the interpretation of a likelihood ratio, whose null χ^2 distribution for nested models would depend on the difference in their degree of freedom p ([37, 39, 40, 41, 42]). The log Bayes factor $2\log B_{12}$ is sometimes called the weight of evidence

Table 1: *Rough Interpretation of Bayes factor B_{12} given by Davison(2003) and Peter(2006)*

B_{12}	$2\log_e B_{12}$	Interpretation
Under 1	Negative	Supports model 1
1 - 3	0 - 2	Weak support for model 2
3 - 20	2 - 6	Support for model 2
20 - 150	6 - 10	Strong evidence for model 2
Over 150	Over 10	Very strong support for model 2

2.3. The Proposed Change Point Model

This section introduces a change-point model based on the Maxwell distribution. Consider a series of observations of size n (n>3) drawn from a Maxwell distribution with parameter θ whose null hypothesis H_1 can be stated as

$$H_1 : \theta_1 = \theta_2 = \theta \tag{7}$$

whose likelihood function can be expressed as

$$f(x \mid \theta) = \prod_{i=1}^{n} \frac{\pi}{2\theta^{2/3}} x^2 e^{-2\theta x^2} \theta > 0 \tag{8}$$

This means that the there is no shift in the parameter θ of the model.
Also, consider a series of observations $x_1, x_2, \ldots x_k, x_{k+1}, \ldots x_n$ with a single shift at point k drawn from different population with parameters θ_1 and θ_2. The alternate hypothesis can be stated as

$$H_2 : \theta_1 \neq \theta_2 \tag{9}$$

having the likelihood function

$$f(x \mid \theta) = \prod_{i=1}^{k} \frac{\pi}{2\theta_1^{2/3}} x^2 e^{-2\theta_1 x^2} \prod_{i=1}^{k} \frac{\pi}{2\theta_2^{2/3}} x^2 e^{-2\theta_2 x^2} \tag{10}$$

2.4. Bayesian Analysis for the Change-Point Model

For the no change-point model (8), we consider a conjugate prior distribution for the parameter θ having $Gamma(\theta|a,b)$ with probability density function and uniform prior for the parameter k with parameter value Uniform(1,n).

$$\pi(\theta) = \frac{b^a}{(n-1)\Gamma(a)}\theta^{a-1}e^{-b\theta} \qquad a > 0, b > 0 \tag{11}$$

The posterior distribution for the null hypothesis model (7) can be obtained by combining the likelihood function (8) with the prior distribution (11) given in Equation (3) as

$$\pi(\theta|x) \propto \frac{\left(b + \frac{\sum_{i=1}^{n} x-i^2}{2}\right)^{\frac{3n}{2}+a}\Gamma(a)}{\Gamma\left(\frac{3n}{2}+a\right)b^a} \tag{12}$$

Also, for the alternate hypothesis model (9), we consider a conjugate prior distribution for θ_1 and θ_2 having $Gamma(\theta_1|a_2,b_2)$ and $Gamma(\theta_2|a_2,b_2)$ having probability density function

$$\pi(\theta_1\theta_2) = \frac{b_1^{a_1}}{\Gamma(a_1)}\theta^{a_1-1}e^{-b_1\theta_1}\frac{b_2^{a_2}}{\Gamma(a_2)}\theta^{a_2-1}e^{-b_2\theta_2} \qquad a_1 > 0, b_1 > 0, a_2 > 0, b_2 > 0 \tag{13}$$

The posterior density function can be derived by combining the likelihood function (10) and prior distribution (13) and thus we have

$$\pi(\theta_1,\theta_2|x) \propto \frac{\left(b_1 + \frac{\sum_{i=1}^{k} x_i^2}{2}\right)^{\frac{3k}{2}+a_1}\Gamma(a_1)\left(b_2 + \frac{\sum_{i=k+1}^{n} x_i^2}{2}\right)^{\frac{3(n-k)}{2}+a_2}\Gamma(a_2)}{\Gamma\left(\frac{3k}{2}+a_1\right)b_1^{a_1}\Gamma\left(\frac{3(n-k)}{2}+a_2\right)b_2^{a_2}} \tag{14}$$

3. Results and Discussions

In this section, we carried out a simulation studies to demonstrate the proposed change-point model. To conduct the diagnostic successfully, we generated five sequences (chains), each consisting of 30,000 elements. A burn-in period of 10,000 observations was implemented, and thinning was applied, considering every 100th observation, using the Markov Chain Monte Carlo (MCMC) scheme.

3.1. Simulation Study

We simulated datasets having a single shift in the parameter θ drawn from a Maxwell distribution with predefined values expressed in model (15)

$$x_i \sim \begin{cases} dMax(1.5) & 1 \le i \le 41 \\ dMax(0.5) & 41 < i \le 80 \end{cases} \tag{15}$$

Table 2: *Summary Statistics for the Posterior Quantities*

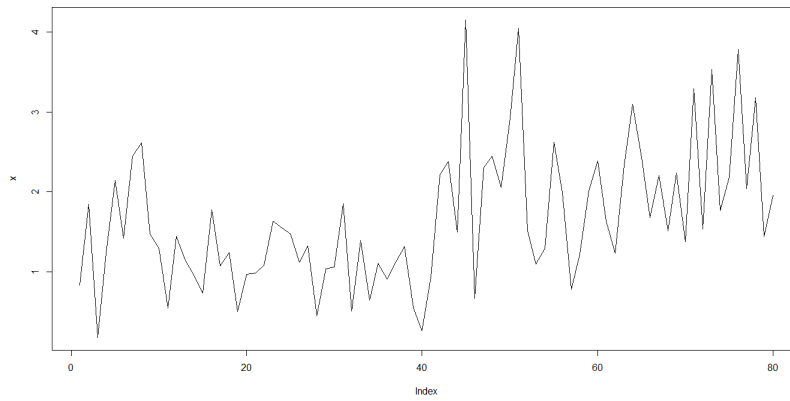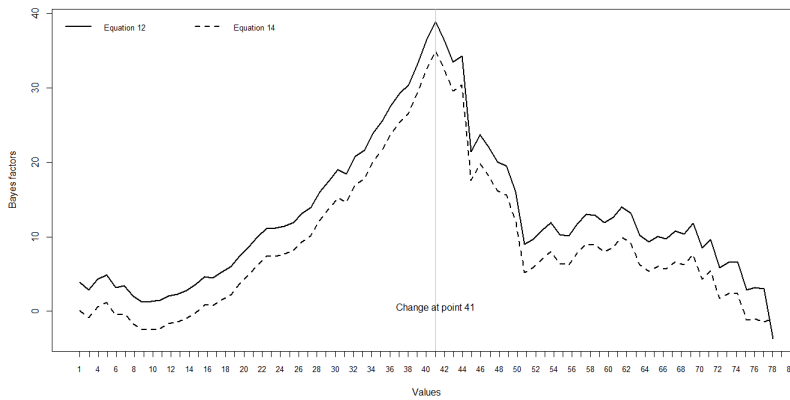| Parameters	Mean	SD	Mc_Error	97.5% Credible Interval
k	41.1	1.137	0.0038	[0.000002–0.0000002]
θ_1	1.77	0.2321	0.0019	[1.348 − 2.255]
θ_2	0.564	0.0743	0.0005	[0.427 − 0.7194]

Figure 1: *Line plot for the simulation study*

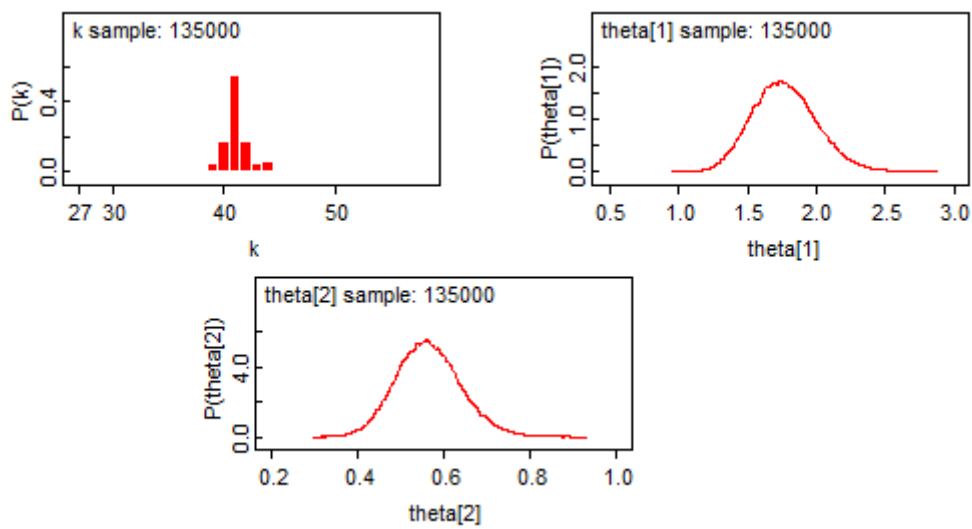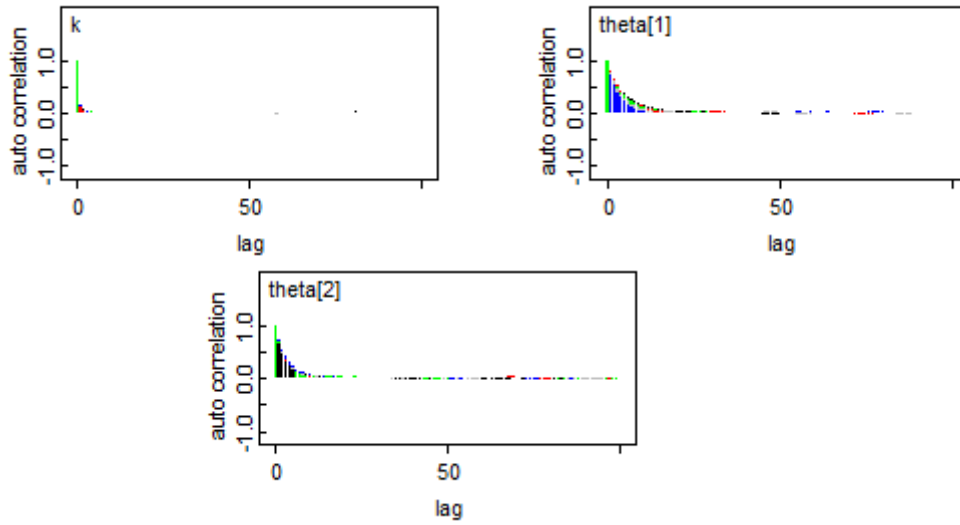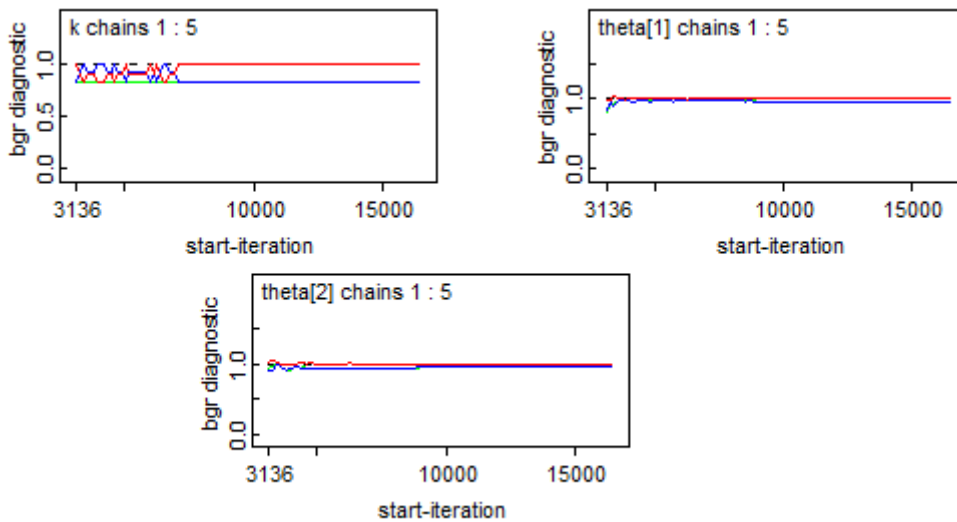**Figure 2:** *Bayes Factor Plot for the simulated dataset*

Figure 3: *Posterior densities for the parameter θ_1, θ_2 and k*

Figure 4: *Autocorrection plot for the Posterior densities for the parameter θ_1, θ_2 and k*

Figure 5: *Brooks-Gelman-Rubin plot for the Posterior densities for the parameter θ_1, θ_2 and k*

Figures 1 and 2 depict the line plot and the Bayes' factor plot for the simulated dataset. From Figure 2, we determined that the shift point was detected at the predefined point 41. Summary statistics from the Gibbs sampling MCMC are shown in Table 2, revealing that the change point k was identified at approximately 41. Posterior densities for all parameters are displayed in Figure 3, confirming that the density of k indicates a change occurring around point 41. The autocorrelation plot in Figure 4 demonstrates noticeable autocorrelation in lag 1 for all parameters. In Figure 5, the Brooks-Gelman plot for the posterior quantities suggests that the chain moves randomly from one iteration to the next, with the Brooks-Gelman-Rubin (BGR) plot for each parameter closely approaching 1. According to Gelman (2003), an acceptable limit is 1.1. Therefore, the BGR plots indicate excellent results. Considering the evidence from Figures 4 and 5, we conclude that the results obtained from the Gibbs sampler exhibit convergence to the posterior distribution and are accurate.

4. Discussion

In this paper, we introduce a single change-point model for datasets that follow a Maxwell distribution, using informative Bayes' Factor techniques. We employ the Bayesian method to detect the time at which a shift occurs in the dataset and apply this approach to both simulated datasets. One key advantage of the Bayesian approach over the frequentist approach is its ability to estimate uncertainty without relying on asymptotic sampling arguments that require large sample sizes. The main objective of this research is to develop a change-point model for detecting a single change-point in a series of observations that follow a Maxwell distribution. We accomplish this using a Bayesian method, which provides a more objective approach compared to subjective methods.

References

[1] Spokoiny, V. Multiscale Local Change Point Detection with Application to Value Art Risk . *Annual Of Statistics*. **37**, 1405-1436 (2009)

[2] Yan, G., Xiao, Z. & Eidenbenz, S. Catching instant messaging worms with change-point detection technique. *In Proceedings Of The USENIX Workshop On Large-scale Exploits And Emergent Threats*. (2008)

[3] Jaxk, R., Cheng, J., Wang, X., Lun, R. & Qiqi, L. A Review and Comparison of Change Point Detection for Climate Data. *Journal Of Applied Meteology And Climatology*. **6** pp. 900-915 (2007)

[4] Londschien, M., Kovacs, S. & Buhlmann, P. Change-point detection for graphical models in the presence of missing values. *Journal Of Computational And Graphical Statistics*. pp. 1-12 (2021)

[5] Erdman, C. & Emerson, J. A Fast Bayesian Change Point Analysis for the Segementation of Micro Arrary Data. *Bioinformatics*. **24**, 2143-2148 (2008)

[6] Kossinets, G. & Watts, D. Empirical analysis of an evolving social network. *Science*. **311** pp. 88-90 (2006)

[7] Eagle, N., Pentland, A. & Lazer, D. Inferring friendship network structure by using mobile phone data. *Proceedings Of The National Academy Of Sciences*. pp. 15274-15278 (2009)

[8] Peel, L. & Clauset, A. Detecting change points in the large-scale structure of evolving networks. *In AAAI*. pp. 2914-2920 (2015)

[9] Killick, R., Eckley, I., Jonathan, P. & Ewan, K. Detection of Changes in the Characteristics of Oceanographic Time-Series Using Statistical Change Point Analysis. *Ocean Engineering*. **37**, 1120-1126 (2010)

[10] Wang, D., Yu, Y., Rinaldo, A. & Willett, R. Localizing changes in high-dimensional vector autoregressive processes. *ArXiv Preprint ArXiv:1909.06359* . (2019)

[11] Safikhani, A. & Shojaie, A. Joint Structural Break Detection and Parameter Estimation in High-Dimensional Nonstationary VAR Models. *Journal Of The American Statistical Association*. **117**, 251-264 (2022)

[12] Isaac, L. & Griffin, L. A historicism in time-series analyses of historical process: Critique, redirection, and illustrations from u.s. labor history. *American Sociological Review*. **54**, 873-890 (1989)

[13] Taylor, C. Change-Point Analysis: A Powerful New Tool For Detecting Changes.. (2010)

[14] Barnett, I. & Onnela, J. Change point detection in correlation networks. *Scientific Reports*. pp. 1-11 (2016)

[15] Zambon, D., Alippi, C. & Livi, L. Change-point methods on a sequence of graphs. *IEEE Transactions On Signal Processing*. **67** pp. 6327-6341 (2019)

[16] Page, E. Continous Inspection Schemes. *Biometrika*. **41** pp. 100-115 (1954,6)

[17] Page, E. A Test for a Change in a Parameter Occuring at an Unknown Point. *Biometrika*. **42** pp. 523-527 (1955)

[18] Page, E. On Problem in which a Change in Parameter Occurs at an Unknown Point. *Biometrika*. **44** pp. 248-252 (1957)

[19] Obisesan, K. Modelling multiple changepoints detection. (University of Ibadan,2015)

[20] Son, Y. & Kim, S. Bayesian single change point detection in a sequence of multivariate normal observations.. *Statistics*. **39**, 373-387 (2005)

[21] Perreault, L., Hache, M., Slivitsky, M. & Bobee, B. Detection of changes in precipitation and runoff over eastern Canada and US using a Bayesian approach. *Stochastic Environmental Research And Risk Assessment*. **13** pp. 201-216 (1999)

[22] Adegoke, T., Bakari, H. & Yahya, A. Bayesian approach for change-point detection of exponential models. *International Journal Of Basic And Applied Sciences*. **17**, 1-6 (2017)

[23] Adegoke, T. & Yahya, ;. A Non-informative Approach to Change-point Detection in a Sequence of Normally Distributed Data with Applications. *Annals Of Statistical Theory And Applications*. **1** pp. 61-70 (2019), www.pssng.org

[24] Adegoke, T. & Yahya, W. A Bayesian Multiple Change-point Analysis: an Application to Air Temperature and Rainfall Data. *Proceedings Of 2nd International Conference Of Professional Statitsical Society Of Nigeria*. **2** pp. 322-327 (2018)

[25] Yahya, W., Obisesan, K. & Adegoke, T. Bayesian Change-point Modelling of Hydrometeorological Data In Nigeria. *Proceedings Of 1st International Conference Of Nigeria Statistical Society*. **1** pp. 59-63 (2017)

[26] Hinkley, D. Inference about the change-point in a sequence of random variables. *Biometrika*. **57** pp. 1-17 (1970)

[27] Fotopoulos, S. & Jandhuala, V. On Hinkley's estimator: Inference about the change point. *Statistics And Probability Letters*., 1449-1458 (2007)

[28] Bisai, D., Chatterjee, S., Khan, A., Trend, N. & Midnapore Weather Observatory, I. Statistical Analysis of Trend and Change Point in Surface Air Temperature Timeseries for Midnapore Weather Observatory, West Bengal, India. *Hydrology Current Research*. **5**, 1-7 (2014)

[29] Liu, P. Maximum Likelihood Estimation of an Unknown Change-Point in the Parameters of a Multivariate Gaussian Series with Applications to Environmental Monitoring. . (Washington State University.,2010)

[30] Dibal, N., Mustapha, M., Adegoke, T. & Yahaya, A. Statistical change point analysis in air temperature and rainfall time series for cocoa research institute of Nigeria, Ibadan, Oyo State, Nigeria. *International Journal Of Applied Mathematics And Theoretical Physics*. **3**, 92-96 (2017)

[31] Ninomiya, Y. Construction of conservative test for change-point problem in two-dimensional random fields. *Journal Of Multivariate Analysis*. pp. 219-242 (2004)

[32] Răkauskas, A. & Suquet, C. Hölder norm test statistics for epidemic change. *Journal Of Statistical Planning And Inference*., 495-520 (2004)

[33] Miller, M. Adapting to Climate Change: Water Management for Urban Resilience. *Environ Urban*. **19** pp. 99-113 (2007)

[34] Grégoire, G. & Hamrouni, Z. Change point estimation by local linear smoothing. *Journal Of Multivariate Analysis*., 56-83 (2002)

[35] Horváth, L. & Kokoszka, P. Change-point detection with non-parametric regression. *Statistics*., 9-31 (2002)

[36] Pandya, M. & Pandya, H. Bayes Estimation of Change Point in Discrete Maxwell Distribution.. *International Journal Of Quality, Statistics, And Reliability*. pp. 1-9 (2011)

[37] Kass, R. & Raftery, A. Bayes Factor. *Journal Of The American Statistical Association*., 773-779 (1995)

[38] Berger, J. & Pericchi The intrinstic Bayes FActor for model selection and prediction. *Journal Of Amer. Stat. Ass.*. **91** pp. 109-202 (1996)

[39] Jeffreys, H. Theory of Probability. (Clarendon Press,1939)

[40] Jeffreys, H. Theory of Probability. (Clarendon Press,1961)

[41] Congdon, P. Bayesian Statistical Modelling. (John Wiley & Sons Ltd,2006)

[42] Davison, A. Statistical Models. (Cambridge University Press,2003)

[43] Gelman, A., and Hill, L.H. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.