Annals of Statistical Theory and Applications (ASTA)
May, 2019, Vol.1: 61-70.

Original Research Article

A Non-informative Approach to Change-point Detection in a Sequence of Normally Distributed Data with Applications

T. M. Adegoke[1]; **W. B. Yahya**[2]

Received: 27th April, 2018 Revised: 4th May, 2019 Accepted: 25th May, 2019

Abstract

In this work, we employed Bayes' Factor approach to detect the point at which change occurred in sequence of normally distributed random variables with an application to data on annual flow of River Nile at Aswan, Egypt. The annual flow of River Nile was modelled by Gaussian density with non-informative prior probability distribution specified for both the mean and variance of the distribution before and after the change-point. The respective posterior distributions were determined via Markov Chain Monte Carlo method based on the simulated and real life data. Various results obtained from simulation studies showed that the Bayesian method developed was capable at detecting the point at which change actually occurred in normally distributed data sets. In the real life data, the method was able to detect the point(period) at which the volume of water in the River Nile changed which was the year 1898 as earlier reported in the literature.

Keywords: Change point analysis, Bayesian method, change in mean level, non-information prior distribution, Gibbs Sampling, Posterior Distribution

1.0 Introduction

Change-point anaysis is a statistial method used to detect if there is a change in the statiscal properties of a statistical distribution. There are a lot of examples where we can detect sudden break in a timeseries dataset such as Nigeria Stock market, climatic variables (such as rainfall, relative humidity, temperature, etc), manufacturing of products (when process is not in control), and so on. Consider the amount of rainfall in a particular region in Nigeria, there is a possibility of observing a time where the annual mean of amount of rainfall for a given year is more/less than the previous years. This changes may also have impact on the variances or not. The change-point model can be linked to the statiscial quality control application, where the manufacturer tends to detect if the process is in control or out of control based the available information.

Change-point models have been applied to different fields which includes Bioinformatics (Erdman and Emerson , 2008), Finance, Fault detection and reliability (Spokoiny, 2009), Signal detection, Surveillance, Security System, Meteorology and Climatology (Jaxk et al., 2007), Environmental Studies (Mohammad-Djafari and Olivier, 2007), Insurance, Econometric Time Series, Macro-sociological process, hydrology (Yahya et al, 2017), Historical changes (Isaac and Griffin, 1989), Detection of Malware software (Yan et

[1] Department of Statistics, University of Ilorin, Ilorin, Nigeria. *E-mail*: adegoketaiwom@hotmail.com
[2] Department of Statistics, University of Ilorin, Ilorin, Nigeria. *E-mail*: wbyahya@unilorin.edu.ng

al., 2008), Oceanography (Killick et al., 2010), Medicine , Finance, Economics (Sangyeol et al., 2006) and Gynecology (Erdman and Emerson, 2008) among others.

Detecting and locating point(s) where change(s) occurred have been under studies for decades under the name "Change-point detection" with a number of methods proposed for detecting abrupt change(s) in a given set of data. These methods comprise of both parametric and non-parametric statistical methods. The change-point analysis started in 1954 when Page published his landmark articles, Page (1954) and Page (1955) using the CUMSUM approach to detect where and when a change occurred in the series of independent normally distributed datasets. Since then, many works on change-point concepts have emerged.

Perreault et al. (2000) developed a single change-point model to detect changes in the series of a multivariate Gaussian data using the Bayesian approach. He assumed informative prior distributions for the parameter of interest and adopted the use of Gibbs sampling MCMC to obtain the estimates of the parameters of interest. Pandya and Pandya (2011) proposed a change point model on discrete Maxwell distribution and detect the point at which the abrupt change occurred using Bayesian methods. The loss functions such as the Squared Error Loss Function, Asymmetric Loss Function, General Entropy Loss Function and Linex Loss Function were used to examine the goodness of the estimated values of the parameters of interest. Nasiri et al. (2014), adopted the use of Bayesian Loss Functions such as Precautionary Loss Function and others to detect change-point and estimate the parameters of interest for any sequence of independent observations that follow Poisson and Geometric distributions.

Son and Seong (2005) developed Bayesian single change-point model in mean and/or covariance of a multivariate Gaussian series by using a non-informative prior distributions. In his work, he employed arithmetic intrinsic Bayes factor, geometric intrinsic Bayes factor and median intrinsic Bayes factor proposed by Berger and Pericchi (1996, 1988) to detect change-point in the data and estimate the parameters of interest.

In this study, the aim is to develop a single change-point model for determining the period of change in the annual flow of water in River Nile using Bayes' factor methodology.

2.0 The Bayes' Factor

Bayesians view hypothesis testing as model comparison (Berger and Pericchi, 1996; Kass and Raftery, 1995). Arguably, the issue is not whether or not a particular hypothesis is true or not, but whether a model described under one hypothesis is preferable to one described under another. The model comparison method was developed by Jeffreys (1939, 1961) to test which of the hypotheses best describe the sequence of an observation. In the study, he obtained a single value known as *Bayes factor*, which is regarded as the ratio of the marginal likelihood of the data in the hypothesis 1(H_1) and hypothesis 2 (H_2). Kass and Raftery (1995) considered observations x with an assumption that it belongs to one of the two hypotheses ω_1 and ω_2 with probability density $\Pr(x/\omega_1)$ or $\Pr(x/\omega_2)$. This can be expressed as

$$H_1: \omega = \omega_1 \text{ vs. } H_2: \omega = \omega_2 \text{ with } \omega_1, \omega_2 \in \omega$$

The prior probabilities are $\Pr(\omega_1)$ and $\Pr(\omega_2) = 1 - \Pr(\omega_1)$. We obtained the posterior distribution for the parameter of intrest given the data for each of the hypotheses denoted as $\Pr(\omega_1/x)$ and $\Pr(\omega_2/x)$ using the Bayes' theorem as,

$$\Pr(\omega_i/x) = \frac{\Pr(x/\omega_i)\Pr(\omega_i)}{\sum_{i=1}^{2}\Pr(x/\omega_i)\Pr(\omega_i)}, \qquad (i = 1,2) \tag{1}$$

However for many applications such as bayesian testing hypothesis on the existence of change-point, it is valuable to use the odds in favour of ω_2 against ω_1 (Davison, 2003; Congdon, 2006) defined by

$$\frac{\Pr(\omega_2/x)}{\Pr(\omega_1/x)} = \frac{\Pr(x/\omega_2)}{\Pr(x/\omega_1)} \times \frac{\Pr(\omega_2)}{\Pr(\omega_1)} \tag{2}$$

which is the multiplication of the prior odds by the bayes factor (BF_{21}). The we have that;

$$BF_{21} = \frac{\Pr(x/\omega_2)}{\Pr(x/\omega_1)} \tag{3}$$

Which implies that;

$$BF_{21} = \frac{likelihood\ of\ observed\ data\ under\ \omega_2}{average\ likelihood\ of\ observed\ data\ under\ \omega_1}$$

The simplest case is when both hypotheses are simple, in which case, BF_{21} equals the likelihood ratio in favour of ω_2. However, both hypotheses involve parameters (say ω_1 and ω_2, the densities $\Pr(x/\omega_i)$, i =1,2, are obtained by integrating over parameter space ω,

$$\Pr(x/\omega_i) = \int \Pr(x/\omega_i)\Pr(\omega_i)\mathrm{d}\omega_i \tag{4}$$

where: ω_i is the parameter under H_i; $\Pr(\omega_i|H_i)$ is the prior density defined on ω_i; $\Pr(x/\omega_i, H_i)$ is the density of y given ω_i and $\Pr(x/\omega_i)$ is the marginal probability distribution

Note that,when computing BF_{21}, all the constants appearing in the definition of a likelihood $(\Pr(y|\omega_i, H_i)$ must be retained. BF_{21} can be expressed in the form

$$BF_{21} = \frac{\Pr(x/\omega_2)}{\Pr(x/\omega_1)} = \frac{\int_{\omega_2}\Pr(x/\omega_2)\Pr(\omega_2)\mathrm{d}\omega_2}{\int_{\omega_1}\Pr(x/\omega_1)\Pr(\omega_1)\mathrm{d}\omega_1} \tag{5}$$

The BF_{21} expressed in (5) is often used to summarize the evidence for ω_2 ω_2 compare to ω_1, with the following heuristic scale below

Table 1: Rough Interpretation of Bayes factor BF_{21} given by Davison (2003) and Congdon (2006)

BF_{21}	$2\log_e BF_{21}$	**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

The BF_{21} is similar to the Likelihood Ratio Test (LRT). The differences beteween BF_{21} and LRT is that the parameters are integrated sinstead of maximizing the parameters under each model. BF_{21} is also known as Bayesian Likelihood Ratio (BLR) which replaces the likelihood with the marginal under both model. This contrasts with the interpretation of a likelihood ratio test whose null χ^2 distribution for nested models would depend on the difference in their degree of freedom (Jeffreys , 1939; 1961; Kass and Raftery, 1995; Davison, 2003; Congdon, 2006). The log Bayes factor $2\log BF_{21}$ is sometimes called the weight of evidence.

3.0 Bayes' Factor of a Normal distribution at an Unknown Change-point τ

Consider the sequence of independent random variables, $x_1, x_2, \ldots, x_n$from a Normal distribution. The objective of this paper is to test the hypotheses of the form

$$H_1: x_i \sim N(\mu, \sigma^2)\ i = 1,2,\cdots,n \tag{6}$$

under the assumption that there is no change-point in mean of a normally distributed random variables, that is $\tau = n$ with $N(\mu, \sigma^2)$ against the alternative that;

$$H_2: x_i \sim \begin{cases} N(\mu_\tau, \sigma^2) & i = 1,2,\cdots,\tau \\ N(\mu_{n-\tau}, \sigma^2) & i = \tau+1, \tau+2, \cdots, n \end{cases} \tag{7}$$

under the hypothesis H_2, we assumed that a change occurred in mean of a normally distributed random variables at point τ^{th} observations, with $\tau < n$, which is the value that indicate the point at which the change occurred. The first τ observations $(x_1, x_2, x_3, \dots, x_\tau)$ come from a Normal distribution with parameters $N(\mu_\tau, \sigma^2)$ and the remaining $n-\tau$ observations $(x_{\tau+1}, x_{\tau+2}, x_{\tau+3}, \dots, x_n)$ come from another Normal distribution with parameters $N(\mu_{n-\tau}, \sigma^2)$. Where $(\mu_\tau, \mu_{n-\tau}) \in \mu$.

$$BF_{21} = \frac{\int\int\int L(\mu_\tau,\mu_{n-\tau},\sigma^2|x)\Pr(\mu_\tau,\mu_{n-\tau})\Pr(\sigma^2)\mathrm{d}\mu_\tau\mathrm{d}\mu_{n-\tau}\mathrm{d}\sigma^2}{\int\int L(\mu,\sigma^2|x)\Pr(\mu)\Pr(\sigma^2)\mathrm{d}\mu\mathrm{d}\sigma^2} \quad (8)$$

where $L(\mu_\tau, \mu_{n-\tau}, \sigma^2|x)$ is the likelihood of the data, $\Pr(\mu_\tau, \mu_{n-\tau})$ is the prior distribution of means and $\Pr(\sigma^2)$ is the prior distribution for the variance under the hypothesis H_1 and $L(\mu, \sigma^2|x)$ is the likelihood function of the data, $\Pr(\mu)$ is the prior probability distribution of means and $\Pr(\sigma^2)$ is the prior probability distribution for the variance under the hypothesis H_0.

The likelihood function for the observations $x_1, x_2, \cdots, x_n$ under the null hypothesis (H_0) can be expressed as

$$Pr(y|\mu, \sigma^2) = \prod_{i=1}^{n} N(\mu, \sigma^2) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2}[nS^2 + n(\bar{x}-\mu)]} \quad (9)$$

where $S^2 = \frac{\sum(x-\bar{x})^2}{n}$ and $\bar{x} = \frac{\sum x}{n}$.

The prior distribution under the null hypothesis H_1 is

$$\Pr(\mu, \sigma^2) = Pr(\sigma^2)Pr(\mu|\sigma^2) = \left(\frac{1}{\sigma^2}\right)\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}} e^{-\frac{1}{2\sigma^2}\mu^2} \quad (10)$$

The posterior distribution under the null hypothesis (H_1) can be obtained by combining (9) and (10)

$$Pr(\mu, \sigma^2|H_1) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2}[nS^2]} e^{-\frac{n+1}{2\sigma^2}(\mu-\psi)^2} \quad (11)$$

The marginal posterior density for the null hypothesis is

$$Pr(x|H_1) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{n+1}\right)^{\frac{1}{2}} \frac{\Gamma(\alpha)}{\beta^\alpha} \quad (12)$$

where $\beta = nS^2$; $\psi = \frac{n\bar{x}^2}{n+1}$; $\alpha = \frac{n}{2}$

Also, for the alternate hypothesis H_2, which claims that there is an existence of a single change-point model by (7), we have that

$$Pr(x|\mu_1, \mu_2, \sigma^2, H_2) = \prod_{i=1}^{\tau} N(\mu_1, \sigma^2) \prod_{i=\tau+1}^{n} N(\mu_2, \sigma^2)$$

$$Pr(x|\mu_\tau, \mu_{n-\tau}, \sigma^2, H_2) = \left(\frac{1}{2\pi\sigma^2}\right)^{\left(\frac{n}{2}\right)} e^{\left\{-\frac{\tau}{2\sigma^2}[S_1^2 + (\bar{x}_1 - \mu_\tau)^2]\right\}} e^{\left\{-\frac{n-\tau}{2\sigma^2}[S_2^1 + (\bar{x}_2 - \mu_{n-\tau})^2]\right\}} \quad (13)$$

where:

$$\bar{x}_1 = \tau^{-1}\sum_{i=1}^{\tau} x; \quad \bar{x}_2 = (n-\tau)^{-1}\sum_{i=(\tau+1)}^{n} x; \quad S_1^2 = \sum_{i=1}^{\tau}\frac{(x-\bar{x}_1)^2}{\tau}; \quad S_2^2 = \sum_{i=\tau+1}^{n}\frac{(x_i-\bar{x}_2)^2}{n-\tau}$$

The prior distribution under the hypothesis H_1 is

$$Pr(\mu_\tau, \mu_{n-\tau}, \sigma^2) = Pr(\sigma^2)Pr(\mu_\tau|\sigma^2)\pi(\mu_{n-\tau}|\sigma^2) = \left(\frac{1}{\sigma^2}\right)\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}} e^{-\frac{1}{2\sigma^2}\mu_\tau^2} \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}} e^{-\frac{1}{2\sigma^2}\mu_{n-\tau}^2} \quad (14)$$

The posterior distribution under the alternate hypothesis (H_2) can be obtained by combining (13) and (14)

$$Pr\left(\mu_{\mu_\tau}, \mu_{n-\tau}, \sigma^2 \middle| H_2\right) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+2} e^{-\frac{1}{2\sigma^2}[\tau S_1^2+(n-\tau)S_2^2]} e^{-\frac{\tau+1}{2\sigma^2}(\mu_\tau-\psi_\tau)^2} e^{-\frac{n-\tau+1}{2\sigma^2}(\mu_{n-\tau}-\psi_{n-\tau})^2} \quad (15)$$

The marginal posterior density is

$$Pr(x|\tau, H_2) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \left(\frac{1}{\tau+1}\right)^{\frac{1}{2}} \left(\frac{1}{n-\tau+1}\right)^{\frac{1}{2}} \frac{\Gamma(\alpha')}{\beta'^{\alpha'}} \quad (16)$$

where $\beta' = \tau S_1^2 + (n-\tau)S_2^2$; $\alpha' = \frac{n}{2} + 2$; $\psi_\tau = \frac{\tau \bar{x}_\tau^2}{\tau+1}$; $\psi_{n-\tau} = \frac{(n-\tau)\bar{x}_{n-\tau}^2}{n-\tau+1}$.

3.1 *Gibbs Sampler: A MCMC Approach*

To obtain the parameters of intrest from the posterior density (16) is not always possible due to the number of parameters involved. So, its quantities is difficult to evaluate analytically (i.e. we may not be able to integrate out some or all of the variables). One of the easiest way to overcome these difficulties in obtaining the estimate of the parameters of intrest is through the Gibbs sampling Markov Chain Monte Carlo (MCMC) method Gelman et al. (2004). In this study, computational techniques based on MCMC algorithms was used to drawout samples from posterior density (16) in order to obtain the estimates of the following parameters $\mu_\tau, \mu_{n-\tau}, \tau$ and σ^2.

The full conditional distribution form all the parameters can be written as:

$$Pr(\mu_\tau|\mu_{n-\tau}, \sigma^2, \tau, H_2) \sim N\left(\mu_\tau|\psi_\tau, \frac{\sigma^2}{\tau+1}\right) \quad (17)$$

$$Pr(\mu_{n-\tau}|\mu_\tau, \sigma^2, \tau, H_2) \sim N\left(\mu_{n-\tau}|\psi_{n-\tau}, \frac{\sigma^2}{n-\tau+1}\right) \quad (18)$$

$$Pr(\sigma^2|\mu_\tau, \mu_{n-\tau}, \tau, H_2) \sim IG(\sigma^2|\alpha', \beta') \quad (19)$$

$$Pr(\tau|H_2) \propto \left(\frac{1}{\tau+1}\right)^{\frac{1}{2}} \left(\frac{1}{n-\tau+1}\right)^{\frac{1}{2}} \frac{\Gamma(\alpha')}{\beta'^{\alpha'}} \quad (20)$$

4.0 Data Analysis

In this section, we provide two applications of Bayesian change-point techniques to determine if indeed a change as occurred or not and at what period. We applied this approach on a simulated dataset and validates its performance on real life dataset. We presented some diagnosis tools to test for convergence of the MCMC generated samples for the obtained parameters. This convergence was verified using Gweke's (1992), Gelman and Rubin's (1992) and Heidelberger and Welsh's (1983) test criterion.

In order to perform the convergence diagnosis, we generated three sequences with 100,000 elements. 10,000 burn-in values were discarded with thinning of 50 observations. The estimate of the parameters $\mu_\tau, \mu_{n-\tau}$, σ^2 and τ are presented in Tables 1 and 2.

4.1 *Analysis of Simulated Data*: *Data Set I*

In order to detect a change-point in a sequence of a normally distributed dataset, 100 random samples were simulated from a Gaussian density as specified by H_2 below with a point of change located at $\left(\frac{2n}{5}\right)^{th}$ observation of the sequence.

$$H_2: x_i \sim \begin{cases} N(14, 7^2) & i = 1,2,\cdots,40 \\ N(46, 7^2) & i = 41,42\cdots,100 \end{cases}$$

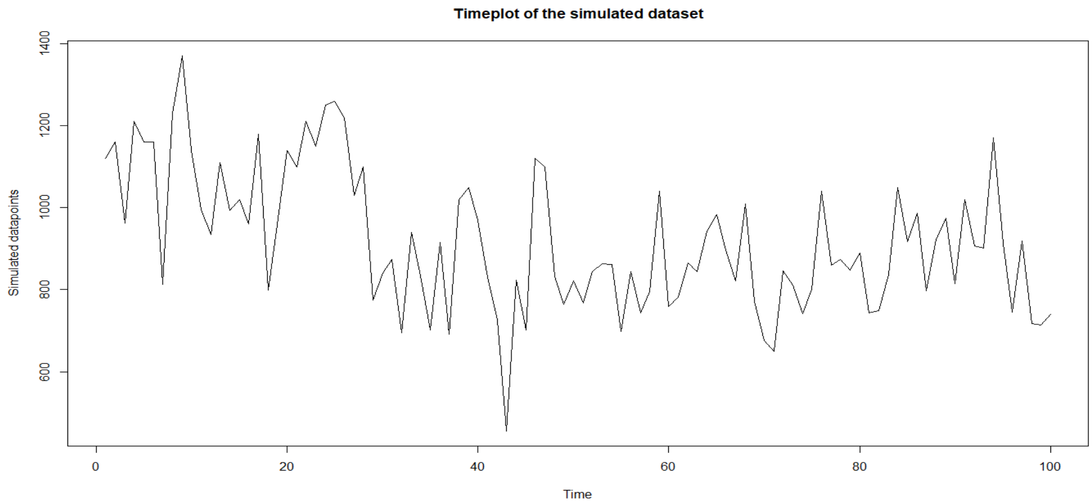

Fig. 1: Timeplot of the simulated dataset

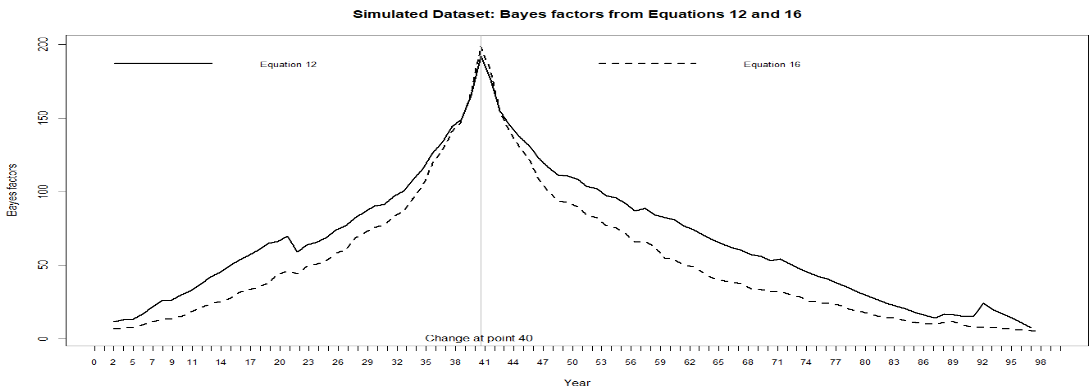

Fig. 2: Bayes' factor plots for the hypotheses H_1 and H_2 for the simulated dataset

Table 1: Posterior quantities for the parameters $\mu_\tau, \mu_{n-\tau}$ and σ^2 for the single point change-point model for the simulated dataset

Parameter	Mean	SD	25%	50%	75%	97.5%
μ_1=μ_τ	13.85	1.08	11.24	13.57	14.25	15.69
$\mu_2 = \mu_{n-\tau}$	46.78	0.91	44.35	46.09	46.72	47.99
σ^2	32.92	1.43	29.77	32.55	33.51	35.55
τ	7.08	0.52	6.09	7.07	7.42	8.17

Table 2: MCMC diagnostics for the simulated data set

Parameter	Max. Autocorrelation Lag 50	Heidelberger and Welsh (Stationary Test)	Gelman and Rubin Shrink Factor	Max. Of the absolute value of Gweke's Criterion
μ_1=μ_τ	0.004	Passed	1.00 ; 1.00	0.06
$\mu_2 = \mu_{n-\tau}$	0.006	Passed	1.00 : 1.01	1.69
σ^2	0.019	Passed	1.01 : 1.05	0.42
τ	0.011	Passed	1.00 : 1.00	0.21

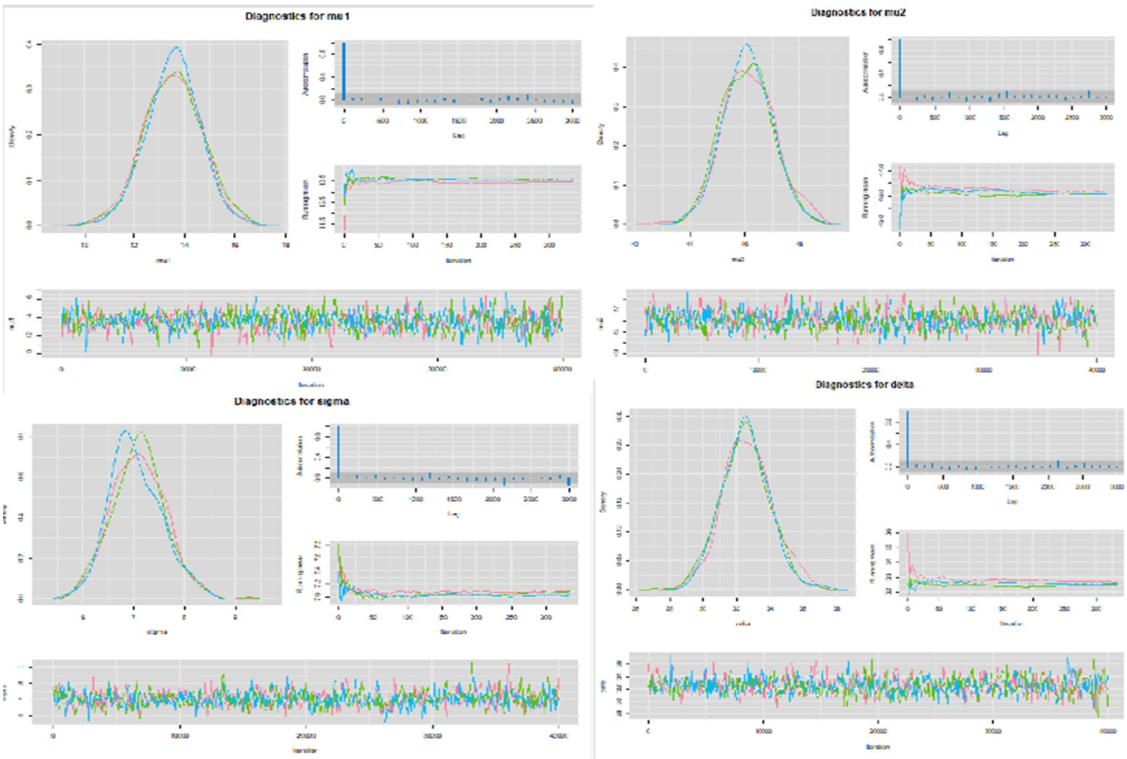

Fig 3: MCMC Diagnostic Plot for simulated dataset

4.2 Analysis of Real life Data: *Data Set II - Annual flow of River Nile*

A single change-point model was developed to detect a change in annual flow of River Nile at Aswan in Egypt from the year 1871to 1980. The data were reported and used by Cobb (1987).

Fig. 4 shows the pattern of the annual flow of River Nile from the year 1871to 1980. Fig. 5 shows the Change-point plot for the annual flow of River Nile which shows that change in the river volume occued and was detected in the year 1898 which literally confirms the claim by Cobb (1978). Fig. 6 provided the MCMC diagnosis of the data.

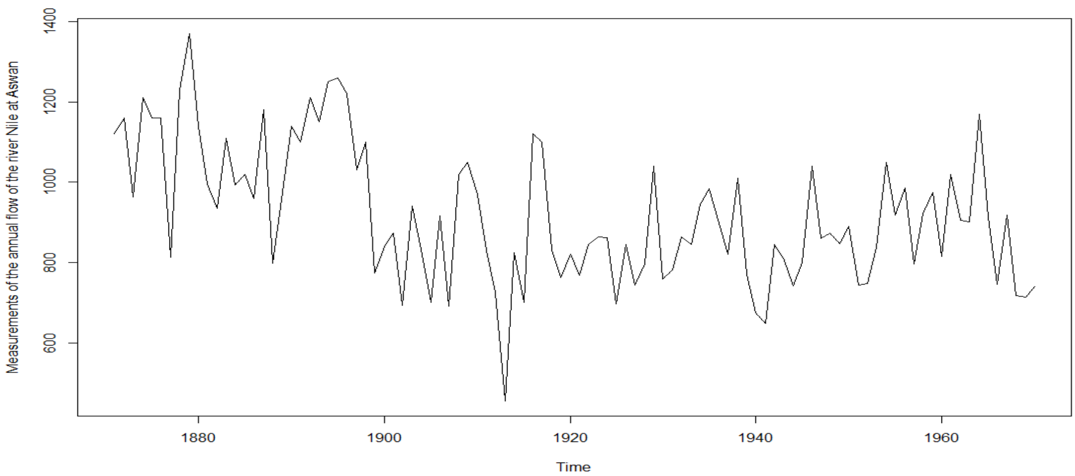

Fig 4: Timeplot of the annual flow of River Nile (Cobb, 1978).

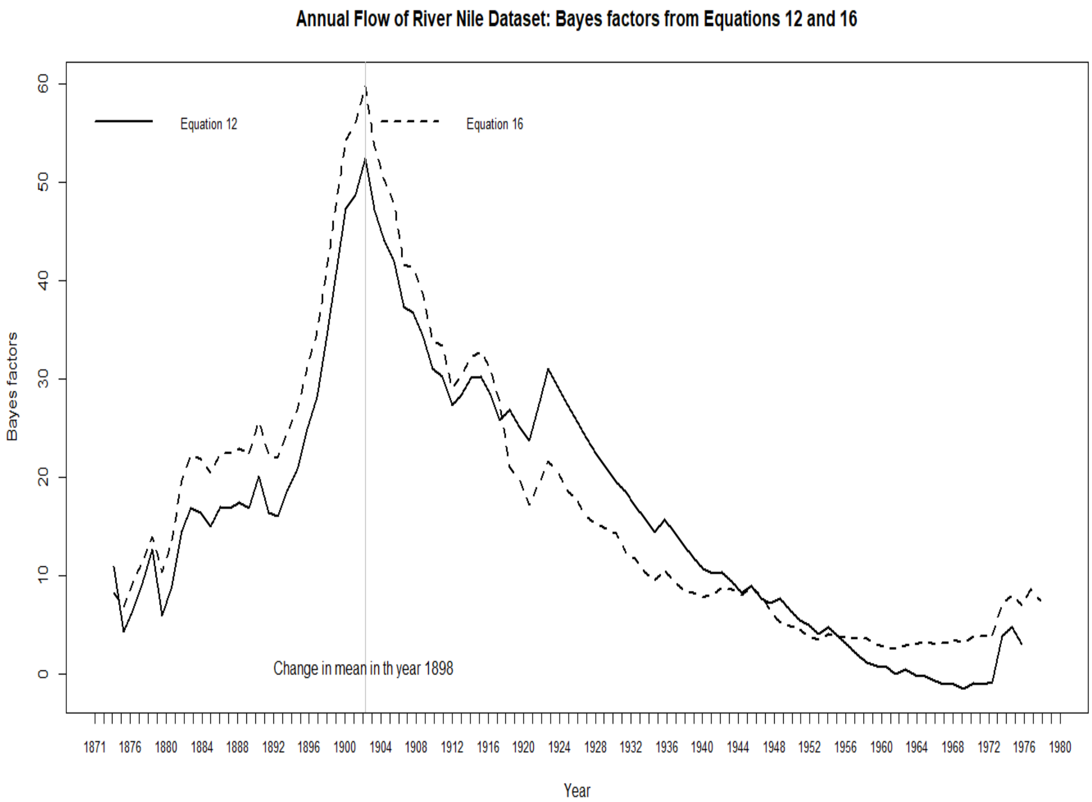

Fig 5: Bayes Factor plot for the annual flow of River Nile. A change was detected in the year 1898.

Table 3: Posterior quantities for the parameters $\mu_\tau, \mu_{n-\tau}$ and σ^2 for the single point change-point model for annual flow of River Nile

Parameter	Mean	SD	25%	50%	75%	97.5%
μ_1=μ_τ	1028.6	26.9	975.3	1028.5	1047.0	1077.0
$\mu_2 = \mu_{n-\tau}$	828.2	15.8	795.0	828.3	839.1	856.6
σ^2	18417	2970	13599.5	17950.0	20182.5	25212.0
τ	-200.4	29.4	29.4	-200.2	-180.6	1282.0

Table 4: MCMC diagnostics for the flow of River Nile

Parameter	Max. Autocorrelation Lag 50	Heidelberger and Welsh (Stationary Test)	Gelman and Rubin Shrink Factor	Max. Of the absolute value of Gweke's Criterion
μ_1=μ_τ	0.050	Passed	1.00 ; 1.02	1.13
$\mu_2 = \mu_{n-\tau}$	0.017	Passed	1.00 ; 1.01	0.76
σ^2	0.010	Passed	1.00 ; 1.02	0.52
τ	0.021	Passed	1.00 ; 1.03	-1.09

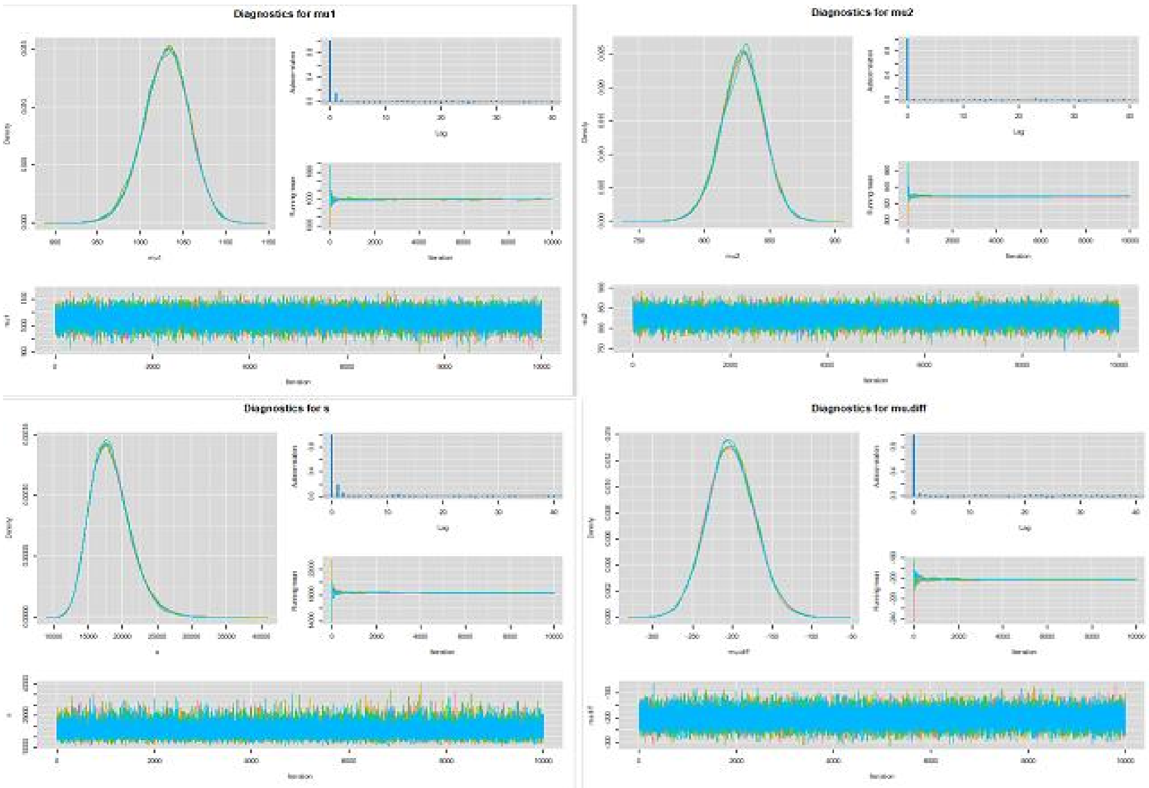

Fig 6: MCMC Diagnostic Plot for flow of River Nile

5.0 Discussion of Results

Figs 2 and 5 are the plots of Bayes' factor against time for both the hypotheses 1 and 2 and we deduced that indeed a change-point was actually detected at point 40 and 28 (1898) for both the simulation and the real life datasets respectively.

As it can be observed from Tables 1 and 3, the autocorrelation is very low for lag size 50. The estimate of the parameters in each of the chains passed the Heidelberger and Welch's stationarity test. Also, for 50% and 97.5% quantiles the Gelman and Rubin's shrink factors for each of the three chains is around 1.00.

Gelman and Rubin (1992) stated that if the shrink factors for 50% and 97.5% quantiles are approximately 1.0, then the posterior quantities converged. The Geweke's maximum z-scores are very moderate which also confirms that the obtained posterior quantities converged.

Figs. 3 and 6 present the graphical summary of the results at different iterations for both simulated and real life data before the change and after the change has occurred at point τ.The parameter traces for each of the generated sequences are displayed below the panels. The obtained posterior density for each of the parameters are shown at the top left hand side panels. Tables 2 and 4 present the summary of quantities that help to describe the features of the posterior distributions of $\mu_{\tau}, \mu_{n-\tau}$ and σ^2 for both the simulation and the real life datasets.

6.0 Conclusions

In this paper, we developed a single change-point model for normally distributed dataset under the non-informative using Baye's Factor techniques. The Bayesian method was used to detect the time at which there is a shift in timeseries dataset and this was employed on both the simulated and real life datasets. The focus of this work is to develop a change-point model for a normally distributed dataset using Bayesian method, rather than a subjective approach.

References

Berger, J. and Pericchi, L. (1988). Accurate and stable bayesian model selection: the median intrinsic bayes factor. *Sankhya Series B*, 60:1–18.

Berger, J. and Pericchi (1996). The intrinstic bayes factor for model selection and prediction. *Journal of American Statistical Association*, 91:109–202.

Cobb, G. W. (1978). The problem of the Nile: conditional solution to a change-point problem. *Biometrika* **65**, 243–51. doi: 10.2307/2335202.

Erdman, C. and Emerson, J. (2008). A fast bayesian change point analysis for the segementation of micro arrary data. *Bioinformatics*, 24(19):2143–2148.

Gelman, A., Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, Hayward, 7(4): 457-511.

Heidelberger, P.; Welch, P. (1983). Simulation run length control in the presence of initial transient. *Operation Research*, Baltimore, 31: 1109-1144.

Isaac, L. W. and Griffin, L. J. (1989). A historicism in time-series analyses of historical process: Critique, redirection, and illustrations from U.S labor history. *American Sociological Review*, 54(6):873–890.

Jaxk, R., Cheng, J., Wang, X., Lun, R., and Qiqi, L. (2007). A review and comparison of change point detection for climate data. *Journal of Applied Meteology and Climatology*, 6:900–915.

Kass, R. and Raftery, A. (1995). Bayes factor. *Journal of the American Statistical Association*, (90):773–779.

Killick, R., Eckley, I., Jonathan, P., and Ewan, K. (2010). Detection of changes in the characteristics of oceanographic time-series using statistical change point analysis. *Ocean Engineering*, 37(13):1120–1126.

Mohammad-Djafari, A. and Olivier, F. (2007). Bayesian approach to change points detection in time series. *International Journal of Imaging Systems and Technoloogy*, 16(5):215–221.

Nasiri, P., Jafari, N., and Jafari, A. (2014). Bayesian estimation of reliability function for a changing exponential family model under different loss functions. *American Journal of Theoretical and Applied Statistics*, 3(1):25–30.

Page, E. (1954). Continous inspection schemes. *Biometrika*, 41:100–115.

Page, E. (1955). A test for a change in a parameter occuring at an unknown point. *Biometrika*, 42:523–527.

Pandya, M. and Pandya, H. (2011). Bayes estimation of change point in discrete maxwell distribution. *International Journal of Quality, Statistics, and Reliability*, pages 1–9.

Perreault, L., Parent, E. ., Bernier, J., Bobee, B., and Slivitsky, M. (2000). Retrospective multivariate bayesian change-point analysis: A simultaneous single change in mean of several hydrological sequences. *Stochasic Environemental Research and Risk Assesment*, 14:243–261.

Sangyeol, L., Nishiyama, Y., and Yoshid, N. (2006). Test for parameter change in diffusion processes by cusum statistics based on one-step estimators. *The Institute of Statistical Mathematics*, 58:211– 222.

Son, Y. and Seong, W. (2005). Bayesian single change point dectection in a sequence of multivariate normal observations. *Statistics*, 5:373–387.

Spokoiny, V. (2009). Multiscale local change point detection with application to value art risk. *Annual of Statistics*, 37(3):1405–1436.

Congdon, P. (2006). *Bayesian Statistical Modelling*. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,West Sussex PO19 8SQ, England.

Davison, A. C. (2003). *Statistical Models*. Cambridge University Press.

Gelman, A.; Carlin, J. B.; Stern, H. S.; Rubin, D. B. (2004). *Bayesian data analysis*. 2nd ed.. Boca Raton: Chapman & Hall.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernardo, J. M.; Berger, J. O.; Dawid, A. P.; Smith, A. F. M. *Bayesian Statistics,* 4. (ed.). Clarendon Press, Oxford, UK.

Jeffreys, H. (1939). *Theory of Probability*. Clarendon Press, Oxford, UK.

Jeffreys, H. (1961). *Theory of Probability*. Clarendon Press, Oxford, UK.

Yahya, W. B., Obisesan. K. O. and Adegoke, T. M. (2017). Bayesian Change-Point Modeling of Rainfall Distributions in Nigeria. *Proceedings of the 1st International Conference of Nigeria Statistical Society,* Vol. 1: 59-63,_edited by Nuamah et al. *www.nss.com.ng*