The Sixteenth International Symposium on Artificial Life and Robotics 2011 (AROB 16th '11),
B-Con Plaza, Beppu,Oita, Japan, January 27-29, 2011

Change Detection Experimentation for Multiple Regression Using ESPRT

--- One Variation is Periodic Function ---

*Katsunori Takeda, *Tetsuo Hattori, **Hiromichi Kawano, ***Tetsuya Izumi

**Graduate School of Engineering, Kagawa University / 2217-20 Hayashi, Takamatsu City, Kagawa 761-0396, Japan*
***NTT Advanced Technology/ Musashino-shi Nakamachi 19-18, Tokyo 180-0006, Japan*
****Micro Technica Co., Ltd. / Yamagami BLD. 3-12-2 Higashi-ikebukuro, Toshima-ku, Tokyo 170-0013, Japan*
(Tel : 81-87-864-2221; Fax : 81-87-864-2262)
(hattori@eng.kagawa-u.ac.jp)

***Abstract*:** Previously, we have proposed the application of sequential probability ratio test (SPRT) to the structural change detection of ongoing time series data. Moreover, we have also proposed the extended method of SPRT (ESPRT). In this paper, we show experimental results by the Extended SPRT (ESPRT) and Chow Test when applying to time series data that are generated by a multiple regression model in the case where one explanatory variation is periodic function (sine function). And we clarify the effectiveness of the ESPRT, in the sense of ability of early and correct change detection and computational complexity.

***Keywords*:** Time series, multiple linear regression, ESPRT (Extended Sequential Probability Ratio Test), Chow Test

I. INTRODUCTION

Generally, we have three stages in predicting ongoing time series data ([1], [2]). First, we have to find a prediction model that adequately represents the characteristics of the early time series data. Second, we have to detect the structural change of the time series data, as quickly and correctly as possible, when the estimated prediction model does not meet the data any more ([3],[4]). Third, we have to reconstruct the next prediction model as soon as possible after the change detection.

For the second problem, we have already proposed an application of SPRT (sequential probability ratio test) that has been mainly used in the field of quality control [5], [6]. And we have presented the experimental results in comparison with Chow Test that is well-known standard method for such structural change detection of time series data ([6], [7]).

However, the experimentation has been done using single regression model and has shown that the SPRT is more effective than Chow Test. Since multiple regression models are more generally used for time series data analysis than single regression one, we have examined by experimentation if the SPRT surpasses Chow Test as well for the case of multiple regression model data [8]. Moreover in the literature [8], we have shown the extended SPRT (ESPRT) aiming at more accurate estimation of the change point.

II. ESPRT AND CHOW TEST

1. SPRT ([8])

The sequential probability ratio test (SPRT) is used for testing a null hypothesis H_0 (e.g. the quality is under pre-specified limit 1%) against hypothesis H_1 (e.g. the quality is over pre-specified limit 1%). And it is defined as follows:

Let $Z_1, Z_2, \cdots Z_i$ be respectively observed time series data at each stage of successive events, the probability ratio λ_i is computed as follows.

$$\lambda_i = \frac{P(Z_1 \mid \mathbf{H}_1) \cdot P(Z_2 \mid \mathbf{H}_1) \cdots P(Z_i \mid \mathbf{H}_1)}{P(Z_1 \mid \mathbf{H}_0) \cdot P(Z_2 \mid \mathbf{H}_0) \cdots P(Z_i \mid \mathbf{H}_0)} \quad (1)$$

where $P(Z \mid H_0)$ denotes the distribution of Z if H_0 is true, and similarly, $P(Z \mid H_1)$ denotes the distribution of Z if H_1 is true.

Two positive constants C_1 and C_2 ($C_1 < C_2$) are chosen. If $C_1 < \lambda_i < C_2$, the experiment is continued by taking an additional observation. If $C_2 < \lambda_i$, the process is terminated with the rejection of H_0 (acceptance of H_1). If $\lambda_i < C_1$, then terminate this process with the acceptance of H_0.

$$C_1 = \frac{\beta}{1-\alpha}, \quad C_2 = \frac{1-\beta}{\alpha} \quad (2)$$

where α means type I error (reject a true null hypothesi s), and β means type II error (accept a null hypothesis as true one when it is actually false).

2. Procedure of SPRT ([8])

The concrete procedure of structural change detection is as follows (see Fig. 1):

Step1: Make a prediction expression and set the tolerance band (***a***) (e.g. ***a***=**2**σ_S) that means permissible error margin between the predicted data and the observed one. (σ_S denotes a standard deviation in learning sample data at early stage.)

Step2 : Set up the null hypothesis H_0 and alternative hypothesis H_1.

H_0 : Change has not occurred yet.

H_1 : Change has occurred.

Set the values α, β and compute C_1 and C_2, according to Equation (2). Initialize i = 0, $\lambda_0 = 1$.

Step3: Incrementing i (i = i+1), observe the following data y_i. Evaluate the error $|\varepsilon_i|$ between the data y_i and the predicted value from the aforementioned prediction expression.

Step4: Judge as to whether the data y_i goes in the tolerance band or not, i.e., the ε_i is less than (or equal to) the permissible error margin or not. If it is Yes, then set $\lambda_i = 1$ and return to Step3. Otherwise, advance to Step5.

Step5: Calculate the probability ratio λ_i , using the following Equation (3) that is equivalent to Equation (1).

$$\lambda_i = \lambda_{i-1} \frac{P(\varepsilon_i \mid \mathbf{H}_1)}{P(\varepsilon_i \mid \mathbf{H}_0)} \tag{3}$$

where, if the data y_i goes in the tolerance band, ($P(\varepsilon_i|H_0)$, $P(\varepsilon_i|H_1)$)= (θ_0, θ_1), otherwise, ($P(\varepsilon_i|H_0)$, $P(\varepsilon_i|H_1)$)= ($(1-\theta_0)$, $(1-\theta_1)$).

Step6: Execution of testing.

(i) If the ratio λ_i is greater than C_2 (= $(1-\beta)/\alpha$), dismiss the null hypothesis H_0, and adopt the alternative hypothesis H_1, and then End.

(ii) Otherwise, if the ratio λ_i is less than C_1 (= $\beta/(1-\alpha)$), adopt the null hypothesis H_0, and dismiss the alternative hypothesis H_1, and then set $\lambda_i = 1$ and return to Step3.

(iii) Otherwise (in the case where $C_1 \le \lambda_i \le C_2$), advance to Step7.

Step7: Observe the following data y_i incrementing i. Evaluate the error $|\varepsilon_i|$ and judge whether the data y_i goes in the tolerance band, or not. Then, return to Step5 (calculation of the ratio λ_i).

Here, we call **Case a-c** for each combination of (θ_0,θ_1), respectively, as follows:

Case a (θ_0=0.1, θ_1=0.9),
Case b (θ_0=0.2, θ_1=0.8),
Case c (θ_0=0.3, θ_1=0.7).

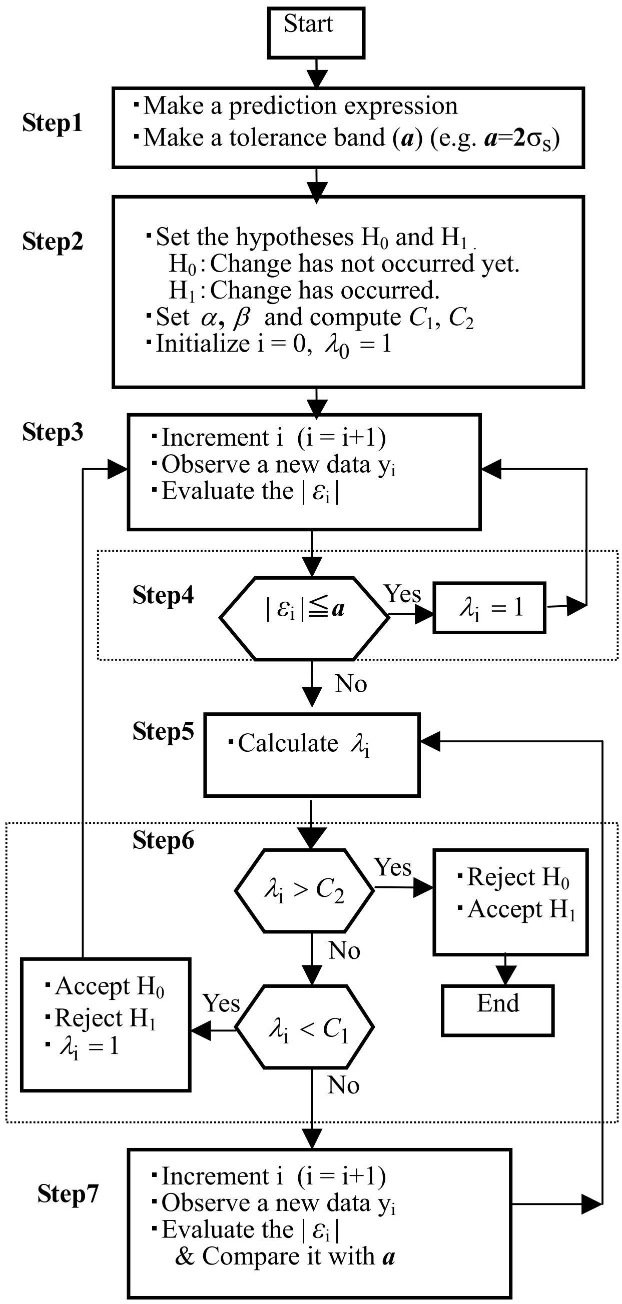

Fig.1. SPRT structural change detection [8].

3. Extended SPRT ([8])

The SPRT detects a change point at the time when the probabilistic ratio λ_i is greater than C_2 (=$(1-\beta)/\alpha$). Then, the detected change point equals to the terminated time point and its detection tends to be delayed from true change point. So, as an extension of SPRT, we define the estimated change point that exists in the intersection [tc-***M***, tc], where tc means the aforementioned detected change point and ***M*** is the number of times when the observed data continuously goes of tolerance zone until the ratio $\lambda_i > C_2$. The number ***M*** can be obtained from the equation (4).

$$\left(\frac{\theta_1}{\theta_0}\right)^M > C_2 \left(= \frac{1-\beta}{\alpha}\right) \qquad (4)$$

Then we have the following equation using Gauss notation. So, the value of ***M*** depends on the parameters (see Table 1). That is, ***M***=2 (Case **a)**, ***M***=3 (Case **b)**, ***M***=4 (Case **c)**.

$$M = \left\lceil \log_{\frac{\theta_1}{\theta_0}} \frac{1-\beta}{\alpha} \right\rceil \qquad (5)$$

Table 1. Parameter values in SPRT and ***M***.

α	β	θ_0	θ_1	***M***
0.05	0.05	0.1	0.9	2
		0.2	0.8	3
		0.3	0.7	4

4. Chow Test ([6]-[8])

The well known Chow Test checks if there are significant differences or not, among residuals for three Regression Lines, where regression Line 1 obtained from the data before a change point tc, Line 2 from the data after tc, and Line 3 from the whole data so far, by setting up the change point hypothesis at each point in the whole data (Fig.2).

III. EXPERIMENTATION

Generally, in the experimentation for the case of time series data based on multiple linear regression model, the data is supposed to be generated by the following equations.

$$y = a_{11}x_1 + a_{12}x_2 + b + \varepsilon \qquad (t \le t_c^*) \qquad (6)$$

$$y = a_{21}x_1 + a_{22}x_2 + b + \varepsilon \qquad (t_c^* \le t) \qquad (7)$$

where $\varepsilon \sim N(0,\sigma^2)$, i.e., the error ε is subject to the Normal Distribution with the average 0 and the variation σ^2, and tc* means the change point. In addition, we have set tc*=70.

Setting the coefficients of equation (6) and (7) as shown in the Table 2, we examine the case where

$$x_1 = t,\ x_2 = \sin\left(\frac{1}{8}\pi t\right) \qquad (8)$$

Table 2. Parameters for generating time series data.

Equation (6) (t=1,2,…,69)	Equation (7) (t=70,71,…,100)	σ
$y = x_1 + 20x_2 + 5 + \varepsilon$	$y = x_1 + 10x_2 + 5 + \varepsilon$	5

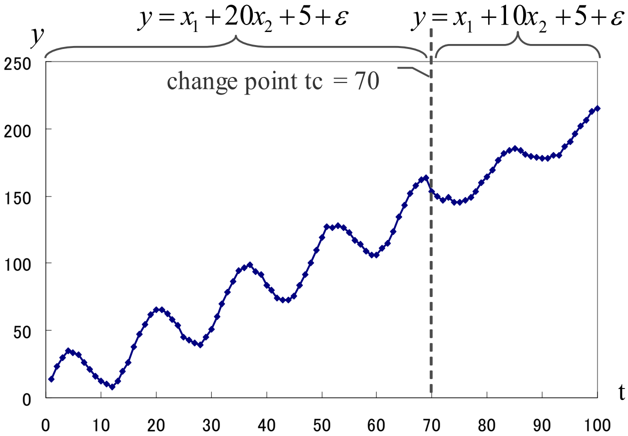

Fig.2. Example of the time series data generated by equations (6) and (7), where x_1 and x_2 are time functions such as $x_1=t$, and $x_2=\sin(\pi t/8)$. (The true change point tc=70.)

Fig.2 shows the example of the time series data according to the equations in Table 2. Fig.3 illustrates results in Chow Test and SPRT, where horizontal axis shows observing time t (detection operation has started from t=41). The vertical axis shows the detected change point tc, whose value is the average of experimentation results for 200 sets of generated time series data.

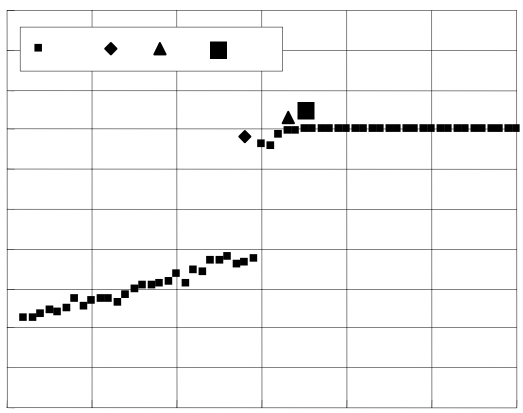

Fig.3. Resultant relation between detected change point tc and time point t, where CT means Chow Test and a-c corresponds to each of Cases a-c, respectively.

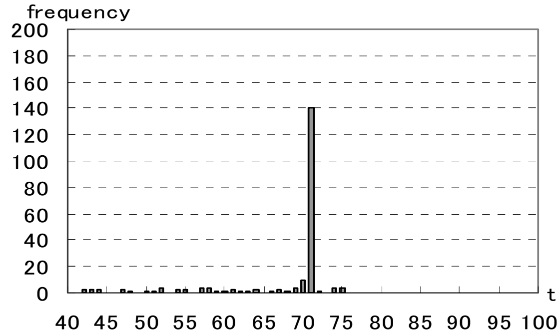

Fig.4. Frequency of the detected change point tc by SPRT in the Case a.

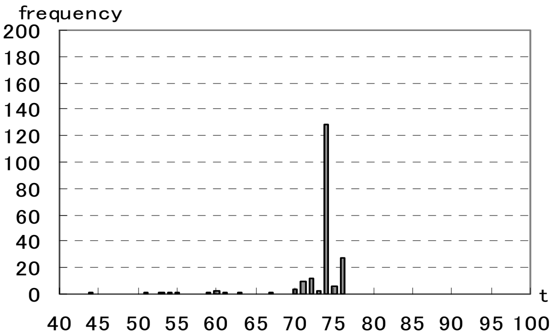

Fig.5. Frequency of the detected change point tc by SPRT in the Case b.

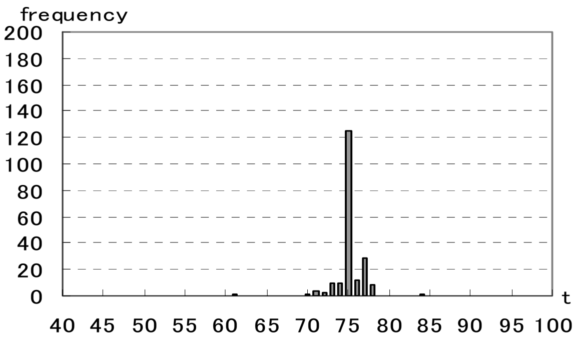

Fig.6. Frequency of the detected change point tc by SPRT in the Case c.

Fig.3 shows that Chow Test outputs the change point at the time when every data is observed after t=40. This means that Chow Test tends to make false detection before the true change point. And, the time when Chow Test works well is long enough after the true change point. Comparing with Chow Test, the SPRT works better depending on the Cases **a**, **b**, **c**. In order to interpret the results as ones by ESPRT, we only have to consider [tc-***M***, tc], where M=2, 3, 4 corresponding to Case a, b, c, respectively. From Fig.4-6, we can see that the ESPRT works better than SPRT and Chow Test.

IV. CONCLUSION

We have experimented the structural change point detection by SPRT, ESPRT, and Chow Test for ongoing time series data generated by multiple linear regression model, where two variations are time functions and the one of two variation is periodic. From the results, we consider that ESPRT works more effective than SPRT and Chow Test in the sense of early detection, accuracy, and computational cost.

REFERENCES

[1] C.G.E.P.Box and G.M.Jenkins (1976), Time Series Analysis:Forecasting and Control, Prentice Hall.
[2] Peter J. Brockwell and Richard A. Davis (2003), Introduction to Time Series and Forecasting, Springer; 2nd edition.
[3] C.Han, P.k.Willet and D.A.Abraham (1999), Some methods to evaluate the performance of Page's test as used to detect transient signals, IEEE Trans. Signal processing, Vol.47, No.8, pp.2112-2127.
[4] S.D.Blostein (1991), Quickest detection of a time-varying change in distribution, IEEE Trans. Information Theory, Vol.37, No.4, pp.1116-1122.
[5] A. Wald (1947), Sequential Analysis, John Wiley & Sons.
[6] Hiromichi Kawano, Tetsuo Hattori, Ken Nishimatsu (2008), Structural Change Point Detection Method of Time Series Using Sequential Probability Ratio Test－Comparison with Chow Test in the ability of early detection－(in Japanese), IEEJ Trans. EIS, Vol.128，No.4，pp.583-592.
[7] Chow, G. C. (1960), Tests of Equality Between Sets of Coefficients in Two Linear Regressions, Econometrica, Vol.28, No.3, pp.591-605.
[8] Katsunori Takeda, Tetsuo Hattori, Izumi Tetsuya, Hiromichi Kawano (2010), Extended SPRT for Structural Change Detection of Time Series Based on Multiple Regression Mode, Proceedings of the 15th International Symposium on Artificial Life and Robotics (AROB 15th '10), pp.755-758, ISBN: 978-4-9902880-4-4, Oita, Japan, Feb..