The Fifteenth International Symposium on Artificial Life and Robotics 2010 (AROB 15th '10),
B-Con Plaza, Beppu,Oita, Japan, February 4-6, 2010

Extended SPRT for Structural Change Detection of Time Series Based on Multiple Regression Model

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

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

Abstract: It is important to quickly detect the structural change of time series as a trigger to remodel the forecasting model. The well-known Chow Test has been used as a standard method for the change detection, especially in economics. On the other hand, we have proposed the application of sequential probability ratio test (SPRT) for the change detection of single regression modeled time series data. In this paper, we show experimental results by SPRT and Chow Test when applying to time series data that are based on multiple regression models. And we clarify the effectiveness of the SPRT comparing with the Chow Test, in the sense of ability of early and correct change detection and computational complexity. Moreover, we extend the definition of the detected structural change point in the SPRT method, and show the improvement of the change detection accuracy.

Keywords: Time series analysis, multiple linear regression, sequential probability ratio test (SPRT), Chow Test

I. INTRODUCTION

There are three types of problems in time series analysis([1],[2]): (i)The first is how to generate a prediction model that adequately represents the characteristics of the early time series data. (ii)The second is how to detect the structural change of the time series, as immediately and correctly as possible, when the estimated prediction model does not meet the real data ([3],[4]), (iii)The third is, after the change detection, how to quickly estimate the time series model again.

For the second problem, we have proposed an application of a sequential probability ratio test (SPRT) ([5],[6])that has been mainly used in the field of quality control, and 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 [7].

Our experimentation has been done just by using single regression model [6] and has shown that the SPRT is more effective than Chow Test. However, multiple regression models are more generally used for time series data analysis than single regression one.

In this paper, we examine by experimentation if the SPRT surpasses Chow Test as well in the change detection of multiple regression model based data. And also, we show the extended SPRT for more accurate estimation of the change point.

II. SPRT AND CHOW TEST

1. SPRT

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

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***=2σ_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).

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

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

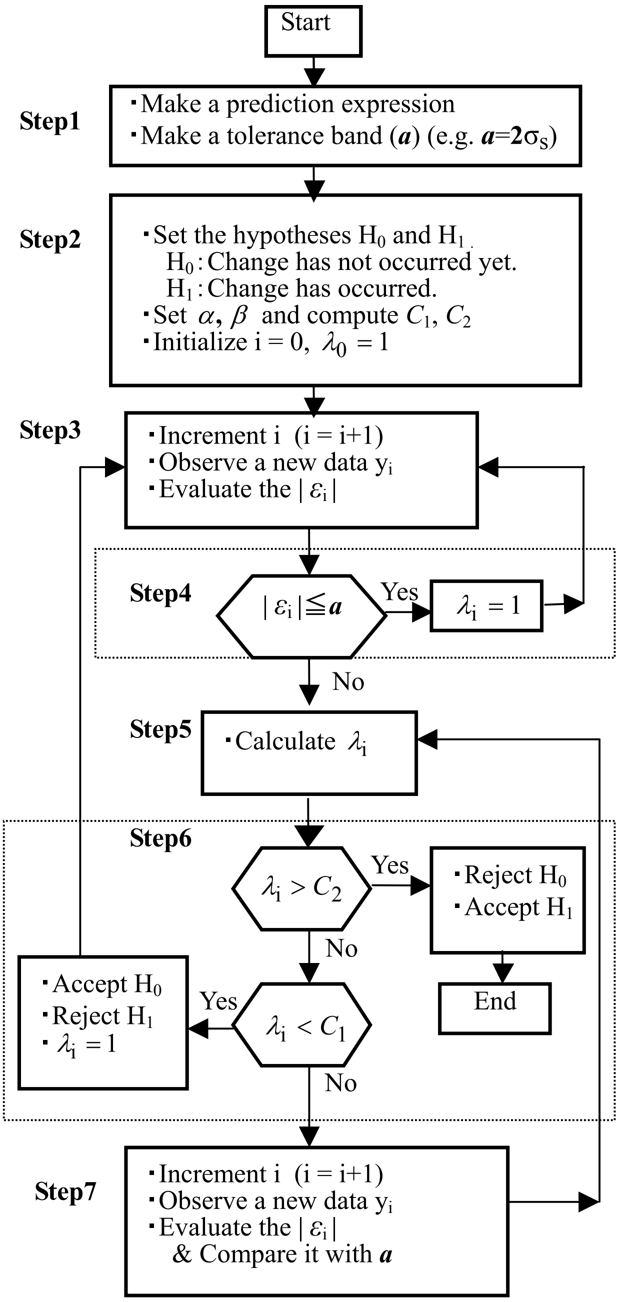

Fig.1. Flow of SPRT structural change point detection.

3. Chow Test

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 hypothesis of change point at each point in the whole data. Fig.2 shows the conceptual image of Chow Test.

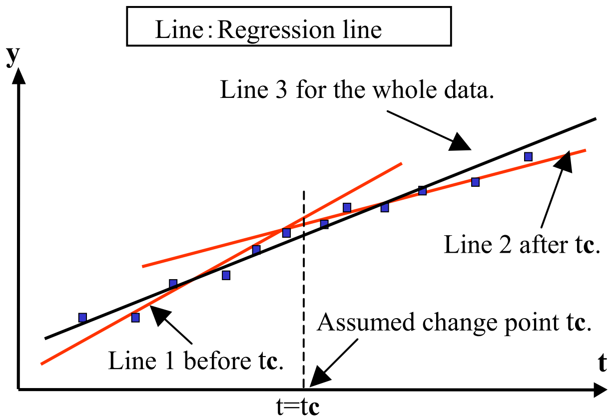

Fig.2. Chow Test in the situation where a hypothesis is set up that the structural change has occurred at t=tc.

III. EXPERIMENTATION

In our experimentation for time series data based on multiple linear regression model, the data is generated by the following equations.

$$y = a_{11}x_1 + a_{12}x_2 + b + \varepsilon \qquad (t \le t_c^*) \qquad (4)$$

$$y = a_{21}x_1 + a_{22}x_2 + b + \varepsilon \qquad (t_c^* \le t) \qquad (5)$$

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.

We have experimented with SPRT and Chow Test for the artificial data based on the above equations (4) and (5). The concrete values of parameters are shown in Table 1. Fig.3 shows an example of the graph of generated time series data by the above equations. Fig.4 shows the situation of change point detections.

The one of the results is illustrated in Fig.5, 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 where the real change point tc*=70.

From Fig.5, we can see that Chow Test outputs the change point at the time when every data is observed after t=40. This means that Chow Test cannot detect the change point properly. And, the time point when Chow Test can work well is long late enough after the real change occurs.

On the other hand, the SPRT detects the change as points cases **a**, **b**, **c**, each corresponding to different parameter values of θ_0, θ_1 that are used in the SPRT.

Table 1. Parameters for generating time series data.

Equation (4) (time t=1,2,…,69)	Equation (5) (time t=70,71,…,100)	σ
$y = 2x_1 + 3x_2 + 10$	$y = 3x_1 + 3x_2 + 10$	5

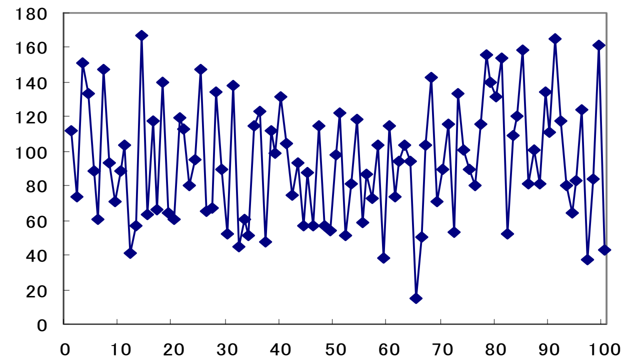

Fig.3. An example of time series data based on multiple linear regression model. Change is occurred at tc=70.

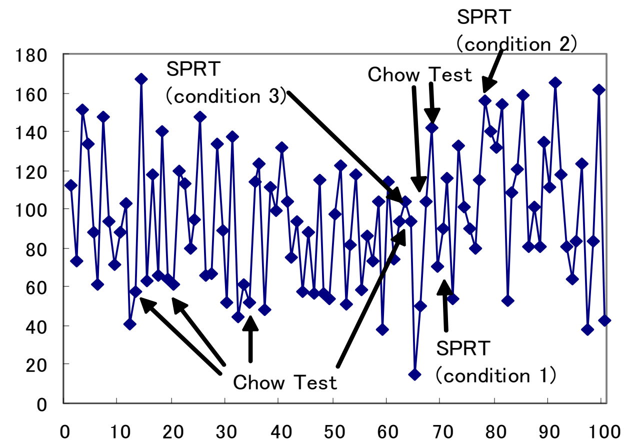

Fig.4. Example of situation of change point detections by SPRT and Chow Test, where tc*=70, and detection operation has started from t=41.

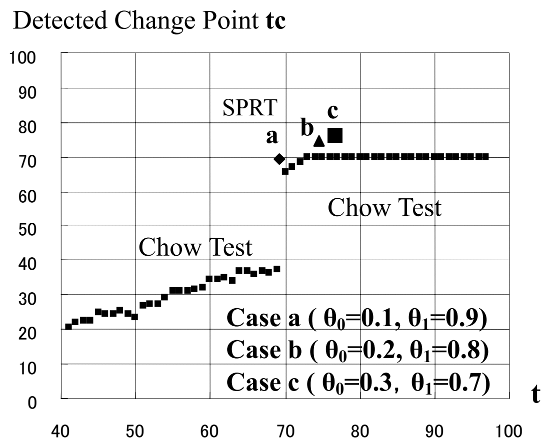

Fig.5. Relation between the observing time t and detected change point tc, where tc*=70, and detection operation has started from t=41.

IV. EXTENDED CHANGE POINT BY SPRT

The SPRT detects a change point at the time when the probabilistic ratio λ_i is greater than C_2 (=(1-β)/α). Then, the detected change point equals to the terminated time point and its detection tends to be delayed from true change point. Thus in this section, we extend the definition of detected change point by SPRT. As such extension, we adopt the number tc-***M*** where tc is ordinary aforementioned 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 (6).

$$\left(\frac{\theta_1}{\theta_0}\right)^M > C_2 \left(= \frac{1-\beta}{\alpha}\right) \quad (6)$$

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

$$M = \left\lceil \log_{\frac{\theta_1}{\theta_0}} \frac{1-\beta}{\alpha} \right\rceil \quad (7)$$

Table 2. 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

Applying the extended definition, we obtain the improvement of hitting percentage in the sense that the detected change point justly equals the true change point. It means pinpoint hitting percentage. Table 3 shows the resultant percentage when using the old definition of detected change point. On the other hand, Table 4 shows the resultant one when using extended definition.

Table 3. Frequency of change point detection and percentage of true change point (TCP) detection.

Case	Time Series Point					Percentage of TCP detection
	70	71	72	73	74	
a	7	141	4	12	1	0.35%
b	1	7	105	5	12	0.05%
c	1	1	5	98	2	0.05%

Table 4. Detected change point frequency and percentage of true change point (TCP) detection.

Case	Time Series Point					Percentage of TCP detection
	70	71	72	73	74	
a	7	141	4	12	1	74.90%
b	1	7	105	5	12	56.50%
c	1	1	5	98	2	50.10%

From those experimental results, we can understand that, if we adopt the interval [tc-***M***, tc] as existing range of true change point, the hitting percentage will considerably increase.

V. CONCLUSION

We have presented and evaluated the experimental results of the structural change detection by SPRT and Chow Test for ongoing time series data that is based on multiple linear regression model. From the results, we consider that SPRT will be more effective in the sense of early detection, accuracy, and computational cost. And also, we are sure that the extended definition of the detected change point is promising and that it would bring about more accurate detection.

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.