Temporal Aggregation Effects on Testing for a Variance Change of a Time Series

Bu Hyoung Lee and William W. S. Wei
Department of Statistics, Temple University, Philadelphia, PA 19122

Abstract
We investigate the effects of temporal aggregation on the cumulative sum of squares (CUSUMSQ) test to detect a variance change in a time series. First, we derive the proper parameter transformation of an aggregate ARIMA model. When temporally aggregated data are used, we show that two aggregation quantities, which are from the aggregate model parameters, in the CUSUMSQ test statistic have effects on test results. Then, we propose a modified CUSUMSQ test to control the aggregation effects. Through Monte Carlo simulations, the modified CUSUMSQ test shows better performance and higher test powers to detect a variance change in an aggregated time series.

Key Words: Temporal Aggregation, Model Transformation, Variance Change, Cumulative Sum of Squares Test

1. Introduction

It is sometimes found that an interruptive event, which starts at a certain time point, causes a structural change in a time series. If there exists a discordance in the second moment before and after the event time, it is called a variance change (or a variance shift) of the series. However, when one is interested to identify the variance change, a statistical test of independent samples, such as the *F*-test, cannot be directly employed because time series observations are almost certainly dependent and no possibility for randomization exists (Box and Tiao, 1965). Also, the event time is often unknown and needs to be estimated; that is, it makes another problem how to distinguish between the pre-event and the post-event and compare their variance difference. Therefore, various alternative procedures to test for a variance change have been proposed and developed in literature. Hsu (1977) proposes a cumulative sum (CUSUM) test for a variance change in a sequence of independent normal random variables when the event time is unknown. Tsay (1988) extends the Hsu test to the cumulative sum of squares (CUSUMSQ) test using time series model errors under independence condition. Inclán and Tiao (1994) show that the limiting null distribution of the CUSUMSQ test statistic follows a Brownian bridge. Lee and Park (2001) and Jin and Zhang (2011) modify the CUSUMSQ test with trimmed observations and bootstrapping, respectively.

Another point of interest is temporal aggregation. Most published time series data are temporally aggregated from the original observations of a small time unit to the cumulative records of a large time unit. However, it is known that temporal aggregation has substantial effects on process properties because it transforms a high frequency nonaggregate series into a low frequency aggregate series. Amemiya and Wu (1972), Brewer (1973), Abraham (1982), Weiss (1984), Stram and Wei (1986), and Silvestrini and Veredas (2008) study the changes of ARIMA model structures and parameters which result from the aggregation. Tiao (1972) and Wei (1978a) show that the aggregate model converges to an IMA limiting model as the aggregation order goes to infinity. Tiao and Wei (1976) and Wei (1978b) discuss the information loss in parameter estimation. Lütkepohl (1984, 1986) investigates the aggregation effects on VARMA model structures and the efficiency of the multivariate forecasts. It is also known that the temporal aggregation strengthens the linearity (Granger and Lee, 1999; Teles and Wei, 2000), induce the normality (Teles and Wei, 2002), and reduce the unit-root characteristic (Teles et al., 2008).

In this paper, we study the effects of temporal aggregation on the CUSUMSQ test to detect a variance change in a time series. The paper is organized as follows. In **Section 2**, we review the test procedure of the CUSUMSQ test. **Section 3** presents aggregation effects on ARIMA model structures and parameters. In **Section 4**, we develop a modified CUSUMSQ test when temporally aggregated data are used. In **Section 5**, we compare the unmodified CUSUMSQ test and the modified CUSUMSQ test through Monte Carlo simulations. Also, some further remarks of the tests are given in **Section 6**.

2. Testing for a Variance Change in a Time Series

The problem of interest is to identify a variance change in a time series process $\{x_t; t = 1,...,n\}$. It can be reworded as testing the null hypothesis of a constant error variance, *i.e.*,

$$H_0 : \sigma^2_{a_1} = \cdots = \sigma^2_{a_n} = \sigma^2_a$$

against the alternative of a variance change starting at a time point k, *i.e.*,

$$H_1 : \sigma^2_{a_1} = \cdots = \sigma^2_{a_{k-1}} \neq \sigma^2_{a_k} = \cdots = \sigma^2_{a_n}$$

for $1 < k \le n$ and $k \in \mathbb{Z}$, where $\mathbb{Z}$ denotes the set of integers and $\sigma^2_{a_t}$ at is an error variance at time t.

We consider two time series processes:

1. A base process $\{x_t^{(0)}; t = 1,...,n\}$, which follows an ARIMA (p,d,q) model of

$$\phi_p(B)(1-B)^d x_t^{(0)} = \theta_q(B)a_t , \qquad (2.1)$$

   where a_t is a Gaussian white noise of mean zero and variance σ^2_a, and $\phi_p(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$ and $\theta_q(B) = 1 - \sum_{j=1}^{q} \theta_j B^j$ are polynomials of backshift operator B. Here, all the roots of $\phi_p(B)$ and $\theta_q(B)$ are assumed to be outside of a unit circle. So the d th difference series $(1-B)^d x_t^{(0)}$ is stationary.
2. A discordant process $\{x_t; t = 1,...,n\}$ with a variance change starting at a time point k, which can be modeled as

$$(1-B)^d x_t = (1-B)^d x_t^{(0)} + v_k \left(\frac{\theta_q(B)}{\phi_p(B)} a_t \right) I_t(k), \tag{2.2}$$

where v_k is a change-magnitude and $I_t(k)=1$ for $t \geq k$ or 0 for $t < k$.

We define an error $e_t = \pi(B)(1-B)^d x_t$ for $t = 1,...,n$, where $\pi(B) = (1-\pi_1 B - \pi_2 B^2 - \cdots)$ $= \phi_p(B)/\theta_q(B)$. From (2.1) and (2.2), we have

$$e_t = \begin{cases} a_t, & \text{for } t < k, \\ (1+v_t)a_t, & \text{for } t \geq k, \end{cases} \tag{2.3}$$

which implies that the error variance changes from σ_a^2 to $(1+v_k)^2 \sigma_a^2$ at k (Tsay, 1988).

Consider the cumulative sum of squares (CUSUMSQ), $\sum_{t=1}^{l} e_t^2$, for $l = 1,...,n$. To test for a variance change at an unknown k, Tsay (1988) and Inclán and Tiao (1994) propose a CUSUMSQ test with the test statistic,

$$\sup_{k=2,...,n} |c_k| \text{ and } c_k = \frac{\sum_{t=1}^{k} e_t^2}{\sum_{t=1}^{n} e_t^2} - \frac{k}{n} . \tag{2.4}$$

Under the null hypothesis of no variance change, it is known that $\sup_{k=2,...,n} |c_k| \xrightarrow{d} \sup_{0<r\leq 1} |B(r)|$ where $B(r) = W(r) - rW(1)$ is a Brownian bridge and $W(r)$ is a Wiener process. Let $\lfloor x \rfloor$ denote the largest integer not greater than a real number x. The value r is chosen from the condition of $\lfloor nr \rfloor = k$ (Inclán and Tiao, 1994).

3. Temporal Aggregation Effects on ARIMA Models

The discordant series x_t is transformed into the m th order temporal aggregate X_T defined to be

$$X_T = \sum_{t=m(T-1)+1}^{mT} x_t = \sum_{j=0}^{m-1} B^j x_{mT} , \tag{3.1}$$

where the aggregation order m is a positive integer for $m < n$ and the aggregate time unit $T = 1,...,N$ for $N = n/m$ (Tiao, 1972; Wei, 2006, p.508). Similarly, the base series $x_t^{(0)}$ is aggregated into $X_T^{(0)} = \sum_{t=m(T-1)+1}^{mT} x_t^{(0)} = \sum_{j=0}^{m-1} B^j x_{mT}^{(0)}$.

It has been known that if $x_t^{(0)}$ follows an ARIMA (p,d,q) model, then its m th order aggregate series $X_T^{(0)}$ is also follows an ARIMA (P,d,Q) model of

$$\Phi_P(\mathcal{B})(1-\mathcal{B})^d X_T^{(0)} = \Theta_Q(\mathcal{B})A_T\,, \tag{3.2}$$

where A_T is a Gaussian white noise of mean zero and variance σ_A^2, and $\Phi_P(\mathcal{B}) = 1-\sum_{i=1}^{P}\Phi_i\mathcal{B}^i$ and $\Theta_Q(\mathcal{B}) = 1-\sum_{j=1}^{Q}\Theta_j\mathcal{B}^j$ are polynomials of backshift operator $\mathcal{B} = B^m$. The orders P and Q are given by

$$P = p\,, \tag{3.3}$$

and

$$Q = \left\lfloor p+d+1-\frac{(p+d+1)-q}{m}\right\rfloor, \tag{3.4}$$

if no hidden periodicity exists in the roots of equations $\phi_p(B)=0$ and $\theta_q(B)=0$. For the details and proofs, we refers readers to Stram and Wei (1986) and Wei (2006, pp.513–515). Also, Amemiya and Wu (1972), Brewer (1973), Weiss (1984), and Silvestrini and Veredas (2008) show aggregate ARIMA model's parameters in terms of nonaggregate ARIMA model's parameters and the aggregation order.

4. Aggregation Effects on the CUSUMSQ Test

Let K be the change point of the discordant series X_T in (3.1), for $1 < K \le N$ and $K \in \mathbb{Z}$. Then, similarly to (2.2), X_T is written as

$$(1-\mathcal{B})^d X_T = (1-\mathcal{B})^d X_T^{(0)} + \mathcal{V}_K\left(\frac{\Theta_Q(\mathcal{B})}{\Phi_P(\mathcal{B})}A_T\right)I_T(K)\,, \tag{4.1}$$

where $(1-\mathcal{B})^d X_T^{(0)}$ is the aggregate stationary series, $\mathcal{V}_K$ is a change-magnitude and $I_T(K)=1$ for $T \ge K$ or 0 for $T < K$.

Similarly to (2.3), an aggregate error E_T is written as

$$E_T = \Pi(\mathcal{B})(1-\mathcal{B})^d X_T = \begin{cases} A_T, & \text{for } T<K, \\ (1+\mathcal{V}_K)A_T, & \text{for } T\ge K, \end{cases} \tag{4.2}$$

which implies that the variance of E_T changes from σ_A^2 to $(1+\mathcal{V}_K)^2\sigma_A^2$ at time point K, where $\Pi(\mathcal{B}) = (q-\Pi_1\mathcal{B}-\Pi_2\mathcal{B}^2-\cdots) = \Phi_P(\mathcal{B})/\Theta_Q(\mathcal{B})$.

In the same manner as (2.4), the CUSUMSQ test statistic for a variance change starting at an unknown K in the series X_T is given by

$$\sup_{K=2,\ldots,N}|C_K| \text{ and } C_K = \frac{\sum_{T=1}^{K}E_T^2}{\sum_{T=1}^{N}E_T^2}-\frac{K}{N}\,, \tag{4.3}$$

of which the null distribution is $\sup_{K=2,\ldots,N} |C_K| \xrightarrow{d} \sup_{0<r\le 1} |B(r)|$.

Let

$$\Delta_T = E_T^2 - \sum_{t=m(T-1)+1}^{mT} e_t^2 . \tag{4.4}$$

Then, C_K in (4.3) is rewritten as

$$C_K = \frac{\sum_{t=1}^{mK} e_t^2 + \sum_{T=1}^{K} \Delta_T}{\sum_{t=1}^{mN} e_t^2 + \sum_{T=1}^{N} \Delta_T} - \frac{K}{N} , \tag{4.5}$$

which implies that the two aggregation quantities $\sum_{T=1}^{K} \Delta_T$ and $\sum_{T=1}^{N} \Delta_T$ explain the aggregation effects on the test statistic.

To control the aggregation effects on the test procedure, we eliminate the two aggregation quantities from C_K and propose a modified CUSUMSQ test.

Definition 4.1. A modified CUSUMSQ test statistic for a variance change of the aggregate series X_T is defined to be

$$\sup_{K=2,\ldots,N} \left|C_K^{(m)}\right| \text{ and } C_K^{(m)} = \frac{\sum_{T=1}^{K} (E_T^2 - \Delta_T)}{\sum_{T=1}^{N} (E_T^2 - \Delta_T)} - \frac{K}{N} . \tag{4.6}$$

5. Simulation Studies of the Aggregation Effects

In this section, we compare the empirical null distributions of the two CUSUMSQ test statistics in (4.3) and (4.6) through Monte Carlo simulations. We also investigate their statistical powers.

5.1. Empirical Null Distributions

Under the null hypothesis of no variance change, we generate 2500 different nonaggregate stationary processes (each size 1800) which follow ARMA (1,1) models, $(1-\phi B)x_t = (1-\theta B)a_t$, for every choice ϕ and $\theta \in \{-0.95, -0.8, -0.5, -0.3, -0.1, 0.1, 0.3, 0.5, 0.8, 0.95\}$, assuming $\phi \ne \theta$ and $a_t \overset{iid}{\sim} N(0,1)$. That is, we simulate 2500×90=225000 nonaggregate stationary series of size 1800. Then we transform the simulated series into their mth order temporal aggregates X_T, as defined in (3.1), for $m = 3, 6, 12, 18, 24, 36$, respectively. We note that each aggregate series X_T follows a stationary ARMA (1,1) process, $(1-\Phi\mathcal{B})X_T = (1-\Theta\mathcal{B})A_T$ and its series size becomes $N = 600, 300, 150, 100, 75, 50$, respectively.

We draw the empirical null distributions of the unmodified test statistic (4.3) and the modified test statistic (4.6) for all the 225000 cases. The expected values of (4.3) are distributed on the left panel and the expected values (4.6) on the right panel of **Figure 5.1**.

We notice the null distribution of (4.3) changes its location rightward and its shape downward as the aggregation order m increases. However, the null distribution of (4.6) is also constant in spite of the aggregation order increment. Hence, we are aware that the distribution location and shape mainly are dependent on the two aggregation quantities $\sum_{T=1}^{K} \Delta_T$ and $\sum_{T=1}^{N} \Delta_T$.

5.2. Test Powers

To examine the statistical powers, we simulate alternative distributions. First, we generate 2500 different nonaggregate stationary processes (each size 1800) which follow the ARMA $(1,1)$ models shown in **Section 5.1**, assuming the change $v_k = 10$ at $k = 901$.

In **Table 5.2**, we present the test powers obtained from the simulations at significance level $\alpha = 0.05$. When comparing their means, the test power using the modified (4.6) is much higher than the test power using the unmodified (4.3) as m increases.

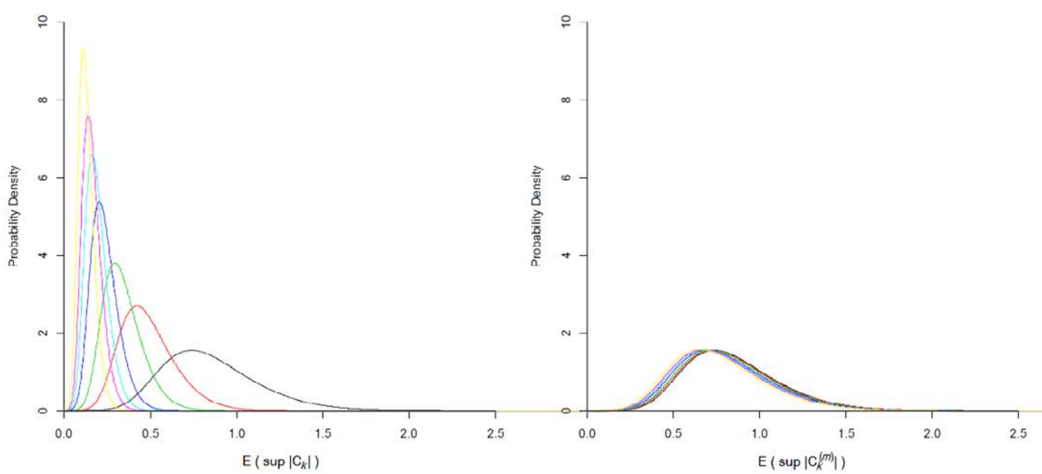

Figure 5.1: Empirical null distributions of the mean CUSUMQ test statistics.

Table 5.2: Test powers of the two CUSUMSQ tests

m	*Mean test powers for*	*Mean test powers for*

	the unmodified test (4.3)	*the modified test* (4.6)
1	0.15904	0.15904
3	0.12316	0.15892
6	0.10764	0.15912
12	0.10524	0.15995
18	0.10547	0.15942
24	0.10543	0.15936
36	0.10297	0.15932

5. Concluding Remarks

In this paper, we analyze the effects of temporal aggregation on the CUSUMSQ test for a variance change in a time series. First, we show the proper model transformation of an ARIMA model structures. Then, using the aggregate model, we derive the modified CUSUMSQ test and find the two aggregation quantities on the test statistic. Through the simulation study and the data examples, we see the modified CUSUMSQ test performs better than the unmodified in terms of test powers. Therefore, we conclude that for efficient variance change detection on aggregate data, we should control the aggregation effect and use the modified CUSUMSQ test.

References

Abraham, B. (1982), "Temporal aggregation and time series," *International Statistical Review*, 50, 285–291.

Amemiya, T., and Wu, R. Y. (1972), "The effect of aggregation on prediction in the autoregressive model," *Journal of the American Statistical Association*, 67, 628–632.

Box, G. E. P., and Tiao, G. C. (1965), "A change in level of a non-stationary time series," *Biometrika*, 52, 181–192.

Brewer, K. R. W. (1973), "Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX models," *Journal of Econometrics*, 1, 133–154.

Granger, C. W. J., and Lee, T. H. (1999), "The effect of aggregation on nonlinearity," *Econometric Reviews*, 18, 259–269.

Hsu, D. A. (1977), "Test for variance shift at an unknown time point," *Journal of the Royal Statistical Society*, *Ser. C*, 26, 279–284.

Inclán, C., and Tiao, G. C. (1994), "Use of cumulative sums of squares for retrospective detection of changes of variance," *Journal of the American Statistical Association*, 89, 913–923.

Jin, H., and Zhang, J. (2011), "Modified tests for variance changes in autoregressive regression," *Mathematics and Computers in Simulation*, 81, 1099–1109.

Lee, S, and Park, S. (2001), "The cusum of squares test for scale changes in infinite order moving average processes," *Scandinavian Journal of Statistics*, 28, 625–644.

Lütkepohl, H. (1984), "Linear aggregation of vector autoregressive moving average processes," *Economics Letters*, 14, 345–350.

Lütkepohl, H. (1986), "Forecasting temporally aggregated vector ARMA processes," *Journal of Forecasting*, 5, 85–95.

Silvestrini, A., and Veredas D. (2008), "Temporal aggregation of univariate and multivariate time series models: a survey," *Journal of Economic Surveys*, 22, 458–497.

Stram, D. O., and Wei, W. W. S. (1986), "Temporal aggregation in the ARIMA process," *Journal of Time Series Analysis*, 7, 279–292.

Tiao, G. C. (1972), "Asymptotic behaviour of temporal aggregates of time series," *Biometrika*, 59, 525–531.

Tiao, G. C., and Wei, W. W. S. (1976). "Effect of temporal aggregation on the dynamic relationship of two time series variables," *Biometrika*, 63, 513–523.

Teles, P., and Wei, W. W. S. (2000), "The effects of temporal aggregation on tests of linearity of a time series," *Computational Statistics and Data Analysis*, 34, 91–103.

Teles, P., and Wei, W. W. S. (2002), "The use of aggregate time series in testing for Gaussianity," *Journal of Time Series Analysis*, 23, 95–116.

Teles, P., Wei, W. W. S, and Hodgess, E. N. (2008), "Testing a unit root based on aggregate time series," *Communications in Statistics - Theory and Methods*, 37, 565–590.

Tsay, R. S. (1988), "Outliers, level shifts, and variance changes in time series," *Journal of Forecasting*, 7, 1–20.

Wei, W. W. S. (1978a), "Some consequences of temporal aggregation in seasonal time series models," *Seasonal Analysis of Economic Time Series* (Ed. A. Zellner), 433–444, U.S. Department of Commerce, Bureau of the Census, Washington, DC.

Wei, W. W. S. (1978b), "The effect of temporal aggregation on parameter," Journal of *Econometrics*, 8, 237–246.

Wei, W. W. S. (2006), *Time Series Analysis: Univariate and Multivariate Methods (2nd ed.)*, Boston, MA: Addison Wesley.

Weiss, A. (1984), "Systematic sampling and temporal aggregation in time series models," *Journal of Econometrics*, 26, 271–281.