PAPER • OPEN ACCESS

Multioutput Least Square SVR Based Multivariate EWMA Control Chart

View the article online for updates and enhancements.

Multioutput Least Square SVR Based Multivariate EWMA Control Chart

Hidayatul Khusna[1], Muhammad Mashuri[1], Suhartono[1], Dedy Dwi Prastyo[1], and Muhammad Ahsan[1]

[1]Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia

m_mashuri@statistika.its.ac.id

Abstract. Multioutput least square SVR has ability to remove serial correlation of process by mapping multivariate input space to multivariate output space. The aim of this research is to propose multioutput least squares SVR based multivariate EWMA control chart to monitor small change of multivariate autocorrelated process. VARMA model with additive and innovative outliers are generated to investigate the performance of proposed control chart. Simulation studies empirically show that multioutput least squares SVR based multivariate EWMA control chart detect either single or consecutive additive outlier takes place at different time in each variable accurately. On the contrary, single innovative outlier in each variable that occurs either at different time or at the same time is detected by multioutput least squares SVR based multivariate EWMA control chart as double out-of control signals.

1. Introduction

Multivariate control chart is one of the most expanding topic in statistical process control [1]. However, many assumptions have to be satisfied to construct multivariate control chart. In most cases, classical control charts assume that variables are independently distributed. Many researchers had investigated the effect of independence assumption violation on control chart performances such as Harris and Ross [2], Johnson and Bagshaw [3], and Noorossana and Vaghefi [4]. Even for small level of autocorrelation, conventional control chart applied to autocorrelated data could increase the average false alarm rate and would decrease the ability to detect changes on a process [5]. Thus, many researchers had proposed multivariate control chart based residual of conventional time series model to monitor multivariate autocorrelated data such as Chan and Li [6], Kalgonda and Kulkarni [7], Theodossiou [8], Kramer and Schmid [9], and Sliwa and Schmid [10].

In practice, conventional time series model usually does not satisfy the underlying assumption. Moreover, it requires predefined structure of the process so that a lot of expertise is needed to apply the model. To overcome this limitation, several researchers suggested the use of Support Vector Regression (SVR) [11,12] as an alternative method. This method can handle both linear and nonlinear time series data. On the other hand, SVR provide global optimal solution that makes it reproducible. Least Square SVR (LS-SVR) is least square version of SVR which replaces quadratic programming problem with linear programming problem [13,14]. Xu et al. [15] combined Multioutput SVR (M-SVR) [16] and multioutput regression [17] as a basic idea to introduce Multioutput LS-SVR (MLS-SVR). Each output variable in MLS-SVR algorithm is permitted to have different slope function.

Hwang [18] developed MLS-SVR based Multivariate Cumulative Sum (MCUSUM) control chart and pointed out that MLS-SVR based MCUSUM chart outperformed LS-SVR based MCUSUM chart. Thus, the aim of this research is to propose MLS-SVR based Multivariate Exponentially Weighted Moving Average (MEWMA) control chart as a solution to monitor multivariate autocorrelated process. Modelling with MLS-SVR is expected to yield global optimum solution, effective computational time, and minimum Mean Square Error (MSE). The performance of proposed control chart is verified using additive and innovative outliers.

2. MLS-SVR Based MEWMA Control Chart

Let $\mathbf{Y}=[y_{tk}]\in R^{n\times\ell}$ is an observable output variable, where $t=1,2,\ldots,n$ define sample size and $k=1,2,\ldots,\ell$ explain number of output variable. Given specific independent and identically distributed samples $\{(\mathbf{x}_1,\mathbf{y}^1),(\mathbf{x}_2,\mathbf{y}^2),\ldots,(\mathbf{x}_n,\mathbf{y}^n)\}$, $\mathbf{x}_t\in R^d$, where d explains the dimension of input variable and $\mathbf{y}^t\in R^\ell$. Given mapping function $\varphi:R^d\to R^h$, assuming all MLS-SVR parameters associated with $\varphi(\mathbf{x})$ so that $\mathbf{w}_t\in R^h(t\in N_\ell)$ can be written as $\mathbf{w}_t=\mathbf{w}_0+\mathbf{v}_t$, where $\mathbf{w}_0\in R^h$, and $\mathbf{v}_t\in R^h(t\in N_\ell)$. Solving parameters $\mathbf{w}_0$, $\mathbf{V}=(\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_\ell)$ and $\mathbf{b}=(b_1,b_2,\ldots,b_\ell)$ simultaneously is analogue to minimize the following objective function [15].

$$\min J(\mathbf{w}_0,\mathbf{V},\mathbf{\Xi})=\frac{1}{2}\left(\mathbf{w}_0^T\mathbf{w}_0\right)+\frac{\gamma''}{2\ell}trace\left(\mathbf{V}^T\mathbf{V}\right)+\frac{\gamma'}{2}trace\left(\mathbf{\Xi}^T\mathbf{\Xi}\right),$$

$$\text{s.t } \mathbf{Y}=\mathbf{Z}^T\mathbf{W}+repmat(\mathbf{b}^T,n,\mathbf{1})+\mathbf{\Xi},$$

where slack variable matrix $\mathbf{\Xi}=(\xi_1,\xi_2,\ldots,\xi_\ell)\in R_+^{n\times\ell}$, $\mathbf{W}=(\mathbf{w}_0+\mathbf{v}_1,\mathbf{w}_0+\mathbf{v}_2,\ldots,\mathbf{w}_0+\mathbf{v}_\ell)\in R^{h\times\ell}$ illustrate MLS-SVR parameter, kernel function matrix $\mathbf{Z}=(\varphi(\mathbf{x}_1),\varphi(\mathbf{x}_2),\ldots,\varphi(\mathbf{x}_n))\in R^{h\times n}$, and regularized parameter $\gamma',\gamma''\in R_+$.

The parameters of MLS-SVR model are estimated with form Lagrange function from the optimization problem, calculate the partial differentiation, and form positive definite linear equation system as follows:

$$\begin{bmatrix} \mathbf{G} & \mathbf{0}_{\ell n\times\ell n} \\ \mathbf{0}_{\ell\times\ell} & \mathbf{M} \end{bmatrix}\begin{bmatrix} \mathbf{b} \\ \mathbf{M}^{-1}\mathbf{N}\mathbf{b}+\boldsymbol{\alpha} \end{bmatrix}=\begin{bmatrix} \mathbf{N}^T\mathbf{M}^{-1}\mathbf{y} \\ \mathbf{y} \end{bmatrix},\tag{1}$$

where, $\mathbf{N}=blockdiag\overbrace{(\mathbf{1}_n,\mathbf{1}_n,\ldots,\mathbf{1}_n)}^{\ell}\in R^{\ell n\times\ell}$ $\mathbf{M}=\mathbf{\Omega}+(\gamma')^{-1}\mathbf{I}_{\ell n}+\left(\ell\big/\gamma''\right)\mathbf{Q}\in R^{\ell n\times\ell n}$, kernel function matrix

$\tilde{\mathbf{K}}=\mathbf{Z}^T\mathbf{Z}\in R^{n\times n}$, $\mathbf{\Omega}=repmat\left(\tilde{\mathbf{K}},\ell,\ell\right)\in R^{\ell n\times\ell n}$, $\mathbf{Q}=blockdiag\overbrace{(\tilde{\mathbf{K}},\tilde{\mathbf{K}},\ldots,\tilde{\mathbf{K}})}^{\ell}\in R^{\ell n\times\ell n}$, positive definite matrix $\mathbf{G}=\mathbf{N}^T\mathbf{M}^{-1}\mathbf{N}\in R^{\ell\times\ell}$, Lagrange multiplier matrix $\boldsymbol{\alpha}=((\alpha_1')^T,(\alpha_2')^T,\ldots,(\alpha_\ell')^T)^T\in R^{\ell n}$, and output variable $\mathbf{y}=(\mathbf{y}_1^T,\mathbf{y}_2^T,\ldots,\mathbf{y}_\ell^T)^T\in R^{\ell n}$. Thus, the solution of linear equation system (1) can be calculated as follows:

1. Solve ϑ and ν from $\mathbf{M}\vartheta=\mathbf{N}$ and $\mathbf{M}\nu=\mathbf{y}$.
2. Calculate $\mathbf{G}=\mathbf{N}^T\vartheta$.
3. Find solution from $\mathbf{b}=\mathbf{G}^{-1}\vartheta^T\mathbf{y}$ and $\boldsymbol{\alpha}=\nu-\vartheta\mathbf{b}$.

Supposed that $\tilde{\boldsymbol{\alpha}}=((\tilde{\alpha}_1')^T,(\tilde{\alpha}_2')^T,\ldots,(\tilde{\alpha}_\ell')^T)^T$ and $\tilde{\mathbf{b}}$ are the solution of linear equation system (1) so the decision function of MLS-SVR can be written as:

$$\hat{f}(\mathbf{x}) = \varphi(\mathbf{x})^T \tilde{\mathbf{W}} + \left(\tilde{\mathbf{b}}\right)^T = \varphi(\mathbf{x})^T repmat\left(\tilde{\mathbf{w}}_0, 1, \ell\right) + \varphi(\mathbf{x})^T \tilde{\mathbf{V}} + \left(\tilde{\mathbf{b}}\right)^T$$

$$= \varphi(\mathbf{x})^T repmat\left(\sum_{k=1}^{\ell} \mathbf{Z}\tilde{\boldsymbol{\alpha}}_k, 1, \ell\right) + \frac{\ell}{\gamma''} \varphi(\mathbf{x})^T \mathbf{Z}\left(\tilde{\mathbf{A}}'\right) + \tilde{\mathbf{b}}^T$$

$$= repmat\left(\sum_{k=1}^{\ell}\sum_{t=1}^{n} \tilde{\alpha}_{tk} K\left(\mathbf{x}, \mathbf{x}_t\right), 1, \ell\right) + \frac{\ell}{\gamma''}\sum_{t=1}^{n} \tilde{\alpha}_t K\left(\mathbf{x}, \mathbf{x}_t\right) + \tilde{\mathbf{b}}^T.$$

Grid search method [19] is used to identify the proper hyper parameter so that minimum MSE is resulted. An evolutionary algorithm [20] can also be employed to optimize SVR parameters.

Supposed that each output variable $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_\ell$ has significant partial autocorrelation function until lag $p_1, p_2, \ldots, p_\ell$ so that input variables of MLS-SVR are selected based on equation $\mathbf{x}_{tk} = (\mathbf{y}_{1,(t-1)}, \ldots, \mathbf{y}_{1,(t-p_1)}, \ldots, \mathbf{y}_{\ell,(t-1)}, \ldots, \mathbf{y}_{\ell,(t-p_\ell)})$. Furthermore, the optimal residuals of MLS-SVR model $e_{tk} = y_{tk} - \hat{f}(\mathbf{x}_{tk})$, $t = 1, 2, \ldots, n$, $k = 1, 2, \ldots, \ell$ are selected using minimum MSE and multivariate normal distribution criteria. These residuals saved in matrix $\mathbf{e}_t$ so that residual based MEWMA [21] statistic is formulated as $\mathbf{Z}_t = \lambda \mathbf{e}_t + (1-\lambda)\mathbf{Z}_{t-1}$, where λ explains smoothing parameter with $0 < \lambda < 1$ and initial value $\mathbf{Z}_0 = \mathbf{0}$. Hence, the statistic of MLS-SVR based MEWMA control chart is calculated as follows:

$$T_t^2 = \mathbf{Z}_t^T \mathbf{\Sigma}_z^{-1} \mathbf{Z}_t, \tag{2}$$

where $\mathbf{\Sigma}_{Z_t} = \frac{\lambda}{(2-\lambda)}\left[1 - (1-\lambda)^{2t}\right]\mathbf{\Sigma}$. A process is said to be in-control if T_t^2 statistics does not exceed the upper control limit H, where the value of H is calculated based on predefined value of ARL0.

3. Result and Discussion

This section is aimed to investigate the performance of MLS-SVR based MEWMA control chart using simulation study. For this purpose, simulation study is designed to generate in-control VARMA model, VARMA model with both additive and innovative outliers. In-control VARMA model with 200 samples are generated using this mathematical expression:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix} + \begin{bmatrix} 0.8 & 0 \\ 0 & 0.9 \end{bmatrix}\begin{bmatrix} y_{1(t-1)} - 5 \\ y_{2(t-1)} - 10 \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} 0.4 & 0 \\ 0 & 0.6 \end{bmatrix}\begin{bmatrix} a_{1(t-1)} \\ a_{2(t-1)} \end{bmatrix},$$

where $\mathbf{a}_t = \begin{bmatrix} a_{1t} & a_{2t} \end{bmatrix}^T$ satisfy white noise residual and bivariate normal distribution with zero mean vector and covariance matrix $\mathbf{\Sigma} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$. These in-control VARMA data, named dataset 1, are shown at Figure 1.
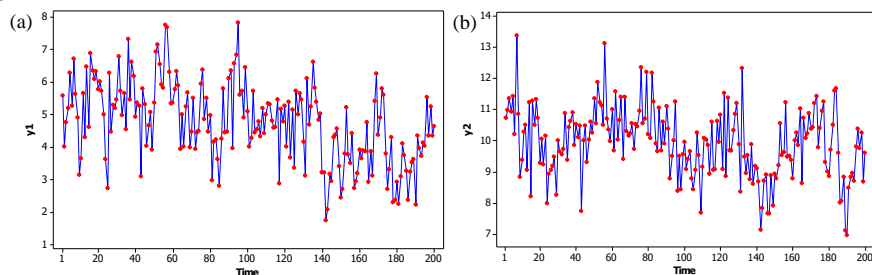

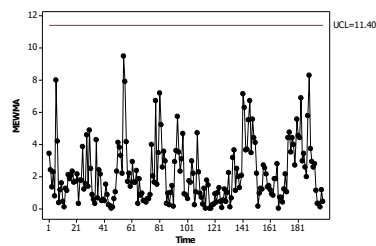**Figure 1**. Time series plot of dataset 1: VARMA model without outlier (a) Y_1, (b) Y_2.

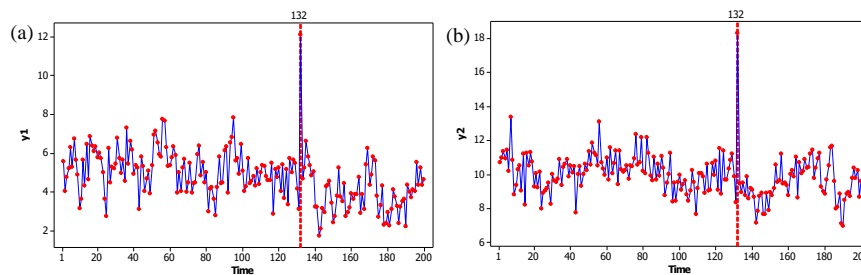**Figure 2**. MLS-SVR based MEWMA control chart for dataset 1.

Figure 3. Time series plot of dataset 2: VARMA with additive outlier at (a) $T_1 = 132$, (b) $T_2 = 132$.
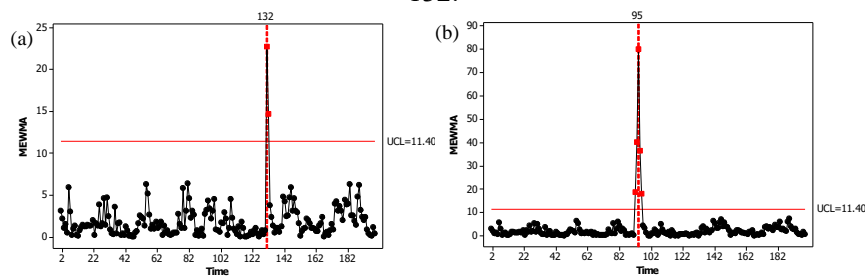

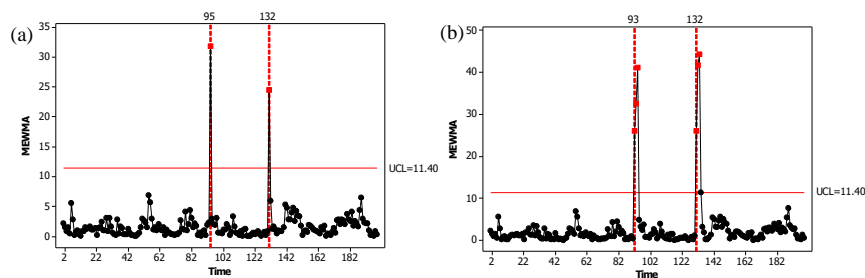**Figure 4.** MLS-SVR based MEWMA control chart for (a) dataset 2, (b) dataset 3.

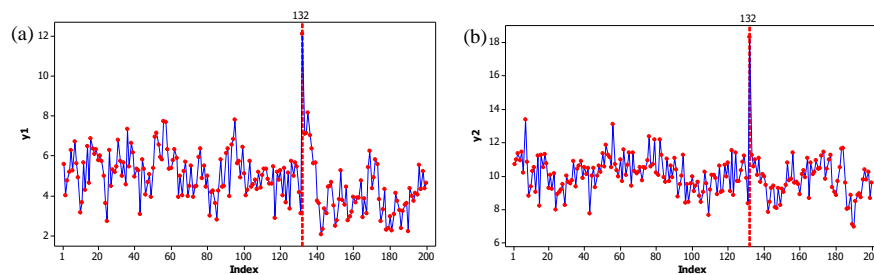**Figure 5.** MLS-SVR based MEWMA control chart for (a) dataset 4, (b) dataset 5.

Figure 6. Time series plot of dataset 6: VARMA with innovative outlier at (a) $T_1 = 132$, (b) $T_2 = 132$.
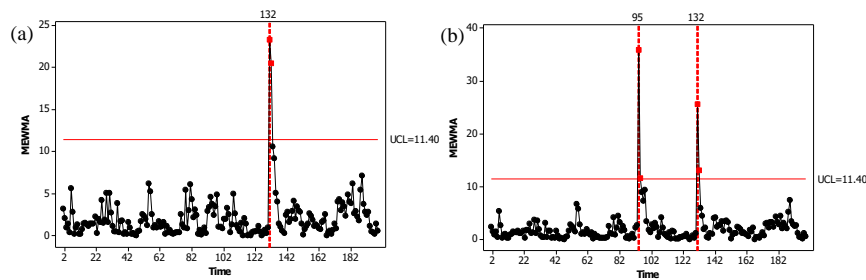
Figure 7. MLS-SVR based MEWMA control chart for (a) dataset 6, (b) dataset 7.

Second simulation is conducted by adding additive outlier at dataset 1 using following mathematical equation:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix} + \begin{bmatrix} 0.8 & 0 \\ 0 & 0.9 \end{bmatrix} \begin{bmatrix} y_{1(t-1)} - 5 \\ y_{2(t-1)} - 10 \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} 0.4 & 0 \\ 0 & 0.6 \end{bmatrix} \begin{bmatrix} a_{1(t-1)} \\ a_{2(t-1)} \end{bmatrix} + 6 \begin{bmatrix} I_{1,t} \\ I_{2,t} \end{bmatrix}; I_{1,t} = I_{2,t} = \begin{cases} 1, & t = 132 \\ 0, & t \neq 132 \end{cases}.$$

Figure 3 displays VARMA model with one additive outlier in each variable at the same time, $T_1, T_2 = 132$, named dataset 2. This outlier only gives effect to the specific time where the outlier actually takes place.

Table 1. Summary of monitoring result

Data	Outlier Type	Monitoring Result
Dataset 1	No outlier	In control
Dataset 2	Additive, $T_1 = T_2 = 132$	OOC at $t = 132,133$
Dataset 3	Additive, $T_1 = T_2 = 93,94,95$	OOC at $t = 93,\dots,97$
Dataset 4	Additive, $T_1 = 95$, $T_2 = 132$	OOC at $t = 95,132$
Dataset 5	Additive, $T_1 = 93,94,95$, $T_2 = 132,133,134$	OOC at $t = 93,94,95$ and $t = 132,133,134$
Dataset 6	Innovative, $T_1 = T_2 = 132$	OOC at $t = 132,133$
Dataset 7	Innovative, $T_1 = 95$, $T_2 = 132$	OOC at $t = 95,96$ and $t = 132,133$

In addition, Table 1 summarizes all datasets resulted in this simulation study.

Dataset 6 are generated by adding innovative outlier at dataset 1. VARMA model with one innovative outlier at each variable, $T_1, T_2 = 132$, can be generated with following equation:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix} + \begin{bmatrix} 0.8 & 0 \\ 0 & 0.9 \end{bmatrix} \begin{bmatrix} y_{1(t-1)} - 5 \\ y_{2(t-1)} - 10 \end{bmatrix} + \left(6 \begin{bmatrix} I_{1,t} \\ I_{2,t} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \right) - \begin{bmatrix} 0.4 & 0 \\ 0 & 0.6 \end{bmatrix} \left(6 \begin{bmatrix} I_{1,t-1} \\ I_{2,t-1} \end{bmatrix} + \begin{bmatrix} a_{1(t-1)} \\ a_{2(t-1)} \end{bmatrix} \right); I_{1,t}, I_{2,t} = \begin{cases} 1, & t = 132 \\ 0, & t \neq 132 \end{cases}.$$

VARMA model with one innovative outlier occurs at the same time in each variable, $T_1, T_2 = 132$, can be shown at Figure 6. This outlier gives effect not only at specific time where the outlier actually takes place but also at some period after the outlier actually takes place, depend on the magnitude of moving average parameter.

Dataset 1 is modelled using MLS-SVR as training data, where $X_{1t} = Y_{1(t-1)}$ and $X_{2t} = Y_{2(t-1)}$ are selected as input variable. MLS-SVR modelling using Radial Basis Function (RBF) kernel function for training data yields the best value of hyper parameter combination $\gamma' = 27$, $\gamma'' = 2\text{-}10$, and $\sigma = 2\text{-}15$ which resulted minimum value of MSE 1.0932. Residuals of MLS-SVR for training data satisfy multivariate normal distribution and white noise condition. Furthermore, these residuals are monitored using MEWMA control chart with significance level $\alpha = 0.00273$ and smoothing parameter $\lambda = 0.3$ as displayed at Figure 2. It can be shown that MLS-SVR based MEWMA control chart is correctly concluded in-control VARMA (1,1) data as an in-control process.

Dataset 2 that displayed at Figure 3 then modelled using MLS-SVR as testing data, while dataset 1 are used as training data. Using the same way as the previous step, the residuals of testing data are monitored using MEWMA control chart. Figure 4.a shows out-of control signal at samples 132 and 133 for an actual additive outlier occurs at the same time in each variable, at $T_1, T_2 = 132$. Furthermore, simulation study using dataset 3 aims to test the performance of MLS-SVR based MEWMA control chart on the presence of additive outliers that occur consecutively in each variable at the same time. Residuals of MLS-SVR for dataset 3 are monitored using MEWMA control chart with the same procedure as previous step as described at Figure 4.b. It can be known that MLS-SVR based MEWMA control chart for dataset 3 points out the out-of control signals at samples 93 until 97 for actual additive outlier that happened in chronological order for each variable at the same time, at $T_1, T_2 = 93,94,95$.

The performance of MLS-SVR based MEWMA control chart in the existence of additive outlier takes place at different time in each variable is tested using dataset 4. Figure 5.a presents MLS-SVR based MEWMA control chart for dataset 4. It can be inferred that MLS-SVR based MEWMA control chart correctly detect single additive outlier takes place at different time in each variable as an out-of control signal. Moreover, dataset 5 performs additive outliers that occur sequentially at different time in each variable, at $T_1 = 93,94,95$ and $T_2 = 132,133,134$. Figure 5.b concludes that MLS-SVR based MEWMA control chart can find precisely the additive outliers that occur consecutively at different time in each variable.

Dataset 6 are generated to investigate the performance of MLS-SVR based MEWMA control chart in the existence of innovative outlier takes place at the same time in each variable. MLS-SVR based MEWMA control chart for dataset 6 is presented at Figure 7.a. It could be noted that MLS-SVR based MEWMA control chart concluded double out-of control signals at samples 132 and 133 for actual single innovative outlier occurs at the same time in each variable, at $T_1, T_2 = 132$. In addition, the performance of MLS-SVR based MEWMA control chart in the presence of innovative outlier takes place at different time in each variable is verified by dataset 7. As provided at Figure 7.b, single innovative outlier occurs at different time in each variable, at $T_1 = 95$ and $T_2 = 132$, are resumed by MLS-SVR based MEWMA control chart as double out-of control signals at samples 95 until 96 and at samples 132 until 133.

Table 1 also summarizes the monitoring result using MLS-SVR based MEWMA control chart. It can be concluded that MLS-SVR based MEWMA control chart yields valid conclusion while monitoring in-control VARMA process and VARMA process with additive outlier, either single outlier or consecutive outlier, takes place at different time in each variable. However, monitoring additive outlier happened at the same time in each variable using MLS-SVR based MEWMA control chart will produce invalid out-of control signal. Single innovative outlier in each variable, either happens at different time or at the same time, also summed up by MLS-SVR based MEWMA control chart as double out-of control signals.

4. Conclusion and Future Research
The simulation study uses VARMA data that involve both additive and innovative outliers. MLS-SVR based MEWMA control chart captures actual additive outlier, either single outlier or consecutive outlier, takes place at different time in each variable as out-of control signal. Meanwhile, additive outlier occurs at the same time in each variable can yield invalid out-of control signal. Moreover, MLS-SVR based MEWMA control chart points out an innovative outlier happens in each variable, either at different time or at the same time, as double out-of control signals. Finally, investigate the effect of autocorrelation in the performance of MLS-SVR based MEWMA control chart using ARL criteria might also be useful for future research.

References
[1]   W.H. Woodall, D.C. Montgomery, J. Qual 1999 *Technol*. **31**, 11.
[2]   T.J. Harris, W H Ross and Can J 1991 *Chem. Eng* **69**, 48.
[3]   R.A. Johnson and M. Bagshaw 1974 *Technometrics* **16**, 103.
[4]   R Noorossana and S J M Vaghefi 2006 *Qual. Reliab. Eng. Int*. **22**, 191.
[5]   S Psarakis and G Papaleonida 2007 *Qual. Technol*. **4**, 501.
[6]   L K Chan and G-Y. Li 1994 *Commun. Stat. - Simul. Comput*. **23**, 997.
[7]   A A. Kalgonda and S R Kulkarni, 2004 *J. Appl. Stat*. **31**, 317.
[8]   P Theodossiou and J Am 1993 *Stat. Assoc*. **88**, 441.
[9]   H G Kramer and L V Schmid 1997 *Seq. Anal*. **16**, 131.
[10]   P Śliwa and W Schmid 2005 *Metrika* **61**, 89.
[11]   U Thissen, R van Brakel, A P de Weijer, W J Melssen, L M C Buydens and Chemom 2003.
        Intell. Lab. Syst. **69**, 35 .
[12]   J R Sato, S Costafreda, P a Morettin and M J Brammer 2008 *Commun. Stat. - Simul. Comput.*
        37, 1183.
[13]   V N Vapnik 1998 *Statistical Learning Theory*.
[14]   V N Vapnik 2008 *The Nature of Statistical Learning Theory*.
[15]   S Xu, X An, X Qiao, L Zhu, L Li 2013 *Pattern Recognit. Lett*. **34**, 1078.
[16]   D Tuia, J Verrelst, L Alonso, F Perez-Cruz and G Camps-Valls 2011 *IEEE Geosci. Remote
        Sens. Lett*. **8**, 804.
[17]   G Liu, Z Lin, Y. Yu 2009 *Pattern Recognit*. **42**, 2737 .
[18]   C Hwang 2016 *J. Korean Data Inf. Sci. Soc*. **27**, 523 .
[19]   C W Hsu, C C Chang and C J Lin, 2016 *Natl. Taiwan Univ*.
[20]   W K   Härdle, D D   Prastyo and C M   Hafner 2014 *Oxford Handb. Appl. Nonparametric
        Semiparametric Econom. Stat*. 346
[21]   C A Lowry, W H Woodall, C W Champ, S E Rigdon, 1992 *Technometrics* **34**, 46.