Association for Information Systems

AIS Electronic Library (AISeL)

ICEB 2005 Proceedings

International Conference on Electronic Business (ICEB)

Winter 12-5-2005

Integrating SPC and EPC for Multivariate Autocorrelated Process

Leeing Tong

Chien-Hui Yang

Cheng-Yi Huang

Cheng-Chi Shou

Integrating SPC and EPC for Multivariate Autocorrelated Process

Lee-Ing Tong[1], Chien-Hui Yang[2], Cheng-Yi Huang[1], Cheng-Chi Shou[1]
[1]Department of Industrial Engineering and Management,
National Chiao Tung University, Hsin-chu, Taiwan 30050, R.O.C
[2]Department of Business Administration,
Yuanpei Institute of Science and Technology, Hsin-chu, Taiwan 30015, R.O.C
EMAIL: litong@cc.nctu.edu.tw
TEL: +886-3-5731896
FAX: +886-3-5733873

Abstract: Statistical process control (SPC) is a widely employed quality control method in industry. SPC is mainly designed for monitoring single quality characteristic. However, as the design of a product/process becomes complex, a process usually has multiple quality characteristics related to it. These characteristics must be monitored by multivariate SPC. When the autocorrelation is present in the process data, the traditional SPC may mislead the results. Hence, the autocorrelated data must be treated to eliminate the autocorrelation effect before employing SPC to detect the assignable causes. Besides, chance causes also have impact on the processes. When the process is out of control but no assignable cause is found, it can be adjusted by employing engineering process control (EPC). However, only using EPC to adjust the process may make inappropriate adjustments due to external disturbances or assignable causes. This study presents an integrated SPC and EPC procedure for multivariate autocorrelated process. The SPC procedure constructs a predicting model using group method of data handling (GMDH), which can transfer the autocorrelated data into uncorrelated data. Then, the Hotelling's T^2 and multivariate cumulative sum control charts are constructed to monitor the process. The EPC procedure constructs a controller utilizing data mining technique to adjust the multiple quality characteristics to their target values. Industry can employ this procedure to monitor and adjust the multivariate autocorrelated process.

Keywords: multivariate process, autocorrelation, statistical process control, engineering process control, group method of data handling, Hotelling's T^2 control chart, multivariate cumulative sum control chart

I. Introduction

Statistical process control (SPC) is a widely employed quality control method in industry. The objective of SPC is to monitor a product/process quality and maintain the process to a fixed target value. The best known tool of SPC is the control chart, which is used to detect the unusual variations in a manufacturing process. When the control chart declares a process out of control, it indicates a problem with the process and the process engineer should look for an assignable cause and try to remove it. Traditionally, control chart is designed for monitoring single quality characteristic. However, as the design of a product/process gets complex, a process usually has multiple related quality characteristics which must be monitored simultaneously. When a process has multiple quality characteristics, it seems reasonable to use a separate univariate control chart for each quality characteristic instead of using a multivariate control chart since the univariate control chart is easily employed and interpreted. However, because each univariate control chart has its own type I error and these univariate control charts are aggregated to monitor multiple quality characteristics simultaneously, the overall probability of a type I error will be increased, in other words, the false alarm will be increased. Another reason that a series of univariate control charts cannot be used to replace the multivariate control chart is these univariate control charts neglect the correlation among the multiple quality characteristics. The consequences of neglecting the correlation among the multiple quality characteristics will increase both type I and type II errors. Therefore, many studies proposed multivariate control charts such as Hotelling's T^2 control chart [4], χ^2 control chart [3], multivariate cumulative sum (MCUSUM) control chart [12,16], and multivariate exponentially weighted moving average (MEWMA) control chart [8] to monitor several quality characteristics simultaneously.

Both univariate and multivariate control charts require that successive measurements from a process are independent of one another. However, due to the effects of tool wear or sampling frequency, the data assembled from a process may exhibit autocorrelation. When the autocorrelation is present in the data, it increases the chance that the control chart will indicate a process shift when the process has not shifted. A few multivariate control charts have been developed to deal with the process with autocorrelated observations. However, these control charts have some practical drawbacks.

The objective of control charts is to detect some unusual variations in the manufacturing process. When a process is out of control but no assignable cause is found, in this case, the process can be adjusted by employing engineering process control (EPC). EPC is a process adjustment technique for control engineers. The concept of EPC is not to remove the assignable causes of departures from the target but to compensate for the drift in process output by

Proceedings of the Fifth International Conference on Electronic Business, Hong Kong, December 5-9, 2005, pp. 692 - 696.

continuously adjusting the process. EPC assumes that there is a dynamic model connecting the process input and the process output. When the model is correct, the EPC technique will minimize the process variability and keep the process output close to its target. However, if assignable causes occur that are outside the framework of a dynamic model, EPC technique cannot compensate the disturbance completely, and consequently, the variability will be increased [11].

Although many studies proposed multivariate control charts to monitor multiple quality characteristics or autocorrelated process, integrating SPC and EPC for a multivariate autocorrelated process is rarely seen. This study presents an integrated SPC and EPC procedure to monitor and adjust multivariate autocorrelated process. The proposed SPC procedure constructs a predicting model using group method of data handling (GMDH) to transfer the autocorrelated data into uncorrelated data. Then, the Hotelling's T^2 and MCUSUM control charts are constructed to monitor the process. The proposed EPC procedure constructs a controller utilizing GMDH to adjust the multiple quality characteristics to their target values. Industry can employ the proposed procedure to monitor and adjust the multivariate autocorrelated process.

II. Literature Review

This section first introduces the multivariate control charts, and then SPC techniques for autocorrelated process are reviewed. In the third subsection, the studies related to EPC are described. The method of GMDH is introduced in the last subsection.

II. 1 Multivariate Control Chart

Hotelling's T^2 control chart was extended from the Shewhart control chart. It was the first developed multivariate control chart for monitoring several identically and independently distributed quality characteristics [4]. Hotelling's T^2 statistic to monitor several quality characteristics simultaneously is employed where assuming the covariance matrix of the quality characteristics is unknown. When the covariance matrix of the quality characteristics is known, the χ^2 statistic is used. Although the computation of Hotelling's T^2 control chart is more complicated than that of Shewhart control chart, when several correlated quality characteristics must be monitored simultaneously, Hotelling's T^2 control chart outperforms Shewhart control chart [4].

The MCUSUM control chart proposed by Woodall and Ncube [16] was extended from the univariate CUSUM control chart. Although the MCUSUM control chart is easy to employ, there may be a delayed reaction to a sudden large shift in the mean. Pignatiello and Runger [12] proposed two types of MCUSUM control charts which were designated as MC1 and MC2 control charts. Their study compared the effectiveness of the two proposed MCUSUM control charts with Hotelling's T^2 control chart and the MCUSUM control chart proposed by Woodall and Ncube. The results indicated that the abilities of detecting small shifts in the mean vector from the MCUSUM control chart proposed by Woodall and Ncube and MC1 control chart are better than that of Hotelling's T^2 control chart. However, when there is a sudden large shift in the mean, the performance of Hotelling's T^2 control chart is better than that of MCUSUM control charts. Furthermore, the average run length (ARL) of MC1 control chart is more stable than the other three control charts.

Hotelling's T^2 control chart only uses the information gained from the current sample to determine whether a process is out of control, therefore, Hotelling's T^2 control chart is insensitive to small shifts in the mean vector. While MCUSUM control chart uses the information gained from the historical and current samples, therefore MCUSUM control chart is sensitive to small shifts in the mean vector.

II. 2 Control Charts for Autocorrelated Process

In recent years, many SPC methods have been developed for monitoring the univariate autocorrelated data, such as the exponentially weighted moving average (EWMA) control chart [6], cumulative sum (CUSUM) control charts [2,13,14], and residual charts [1,9]. A few multivariate SPC methods for monitoring the autocorrelated process are also developed. Notably, Theodossiou [15] has proposed a CUSUM chart and Kramer and Schmid [7] have proposed a multivariate EWMA control chart to monitor the autocorrelated process. However, there are some practical drawbacks of using these methods. A prominent drawback is when evidence of an out-of-control situation is observed, and the above methods fail to provide the quality characteristic(s) responsible for the data point which falls outside the control limits.

II. 3 Engineering Process Control

In mechanical, chemical, and electrical engineering applications, EPC are usually implemented to reduce process's variability. EPC assumes that a variable, which is called manipulatable variable, can be adjusted to compensate for the drift in process output to keep the process output close to the target. EPC assumes that there is a dynamic model connecting the process input and the process output. If the model is correct, the EPC technique will minimize the process variability and keep the process output close to the target. However, if the assignable causes occur that are outside the framework of a dynamic model, the EPC technique cannot compensate the disturbance completely; in that case, the variability will be increased [11]. Montgomery *et al.* [10] have demonstrated that the performance of an integrated SPC and EPC procedure is more effective than the EPC procedure alone.

II. 4 Group Method of Data Handling

Group Method of Data Handling (GMDH) [5] is applied in a great variety of areas in data mining. Inductive GMDH algorithms aim to find interrelations of variables in a data set and select an optimal network or model structure. GMDH is an iterative method which successively tests models selected from a set of candidate models according to a specified criterion. A general connection between the input and output variables can be found in the form of a functional Volterra series. A discrete analogue of Volterra series is Kolmogorov-Gabor polynomial which can be expressed as follows:

$$y = a_0 + \sum_{i=1}^{M} a_i x_i + \sum_{i=1}^{M}\sum_{j=1}^{M} a_{ij} x_i x_j + \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M} a_{ij} x_i x_j x_k + \ldots \quad (1)$$

where $X = (x_1, x_2, \ldots, x_M)$ is the vector of input variables and $A = (a_1, a_2, \ldots, a_M)$ is the vector of summand coefficients.

The combinational GMDH algorithm has a multilayer iterative structure. The specific feature of GMDH is that the iteration rule does not remain consistent but expands with every new series. In the first series, all the models of the simplest structure are in the following form:

$$y = a_0 + a_i x_i \quad i = 1,2,\ldots,M \quad (2)$$

After sorting these models, select the best F models by a specified criterion.

In the second series, models of more complex structure are sorted. These models are constructed on output variables from the best models of the first series. The form of the models of the second series can be expressed as follows:

$$y = a_0 + a_1 x_i + a_2 x_j$$
$$i = 1,2,\ldots,F;\ j = 1,2,\ldots,M.\ F \le M \quad (3)$$

In the third series, the sorting involves more complex structure form as follows:

$$y = a_0 + a_1 x_i + a_2 x_j + a_3 x_k$$
$$i = 1,\ldots,F;\ j = 1,\ldots,F.\ \mathrm{k} = 1,\ldots,\mathrm{M} \quad F \le M \quad (4)$$

The iterative procedure of the series continues until the criterion value stops increasing.

III. Proposed Procedure

This study develops an efficient procedure which includes two stages for monitoring a multivariate autocorrelated process. In the first stage, the SPC procedure constructs a predicting model using GMDH to transfer the autocorrelated data into uncorrelated data. Then, the Hotelling's T^2 and MC1 control charts are constructed to monitor a process to detect both large and small shift in the mean at the same time. If the process is out of control but no assignable cause is found, it can be adjusted by employing EPC to keep the multiple quality characteristics close to their target values. Therefore, in the second stage, the EPC procedure is proposed to construct a controller utilizing GMDH to adjust the multiple quality characteristics to their target values. The proposed two-stage procedure is described in Section 3.1 and 3.2.

III. 1 The Statistical Process Control Procedure

Constructing a multivariate control chart requires the analysis of a preliminary data set that is assumed to be in control and this data set is then used to construct a control chart for monitoring the process. Therefore, this section includes two phases. The first phase, designated as phase I, is to estimate parameters that will be used subsequently for on-going monitoring of the process. Phase I should utilize a very large sample of data so that the parameters and control limits are well-estimated for Phase II. Then, the second phase, designated as phase II, is to use the constructed control chart to monitor the process. The two phases are described as follows:

III. 1. 1 Phase I of the Statistical Process Control Procedure

This subsection presents a procedure to construct multivariate control charts for autocorrelated process. The procedure is described as follows:

Step 1: Collect and analyze the process data.

Assume that there are *i* input variables $x_1, x_2 \ldots x_i$ and *j* quality characteristics $y_1, y_2 \ldots y_j$ in the process. Draw univariate Shewhart control charts for the *j* quality characteristics to gain the information of the process data.

Step 2: Calculate the sample autocorrelation function of the data of each quality characteristic

Calculate the sample autocorrelation function for each quality characteristic and draw an autocorrelation function plot for each quality characteristic. If there is at least one quality characteristic exhibiting autocorrelation, go to Step 3; otherwise, go to Step 5.

Step 3: Construct a predicting model using GMDH

In order to eliminate the autocorrelation presented in the process data, this study utilizes GMDH as a predicting model to transfer the autocorrelated data into uncorrelated data to satisfy the assumption of independence of control charts. This study uses the observations of *N* periods before time *t* from the autocorrelated quality characteristic as the process input, and the process output is the quality characteristic of time *t*. After constructing the GMDH predicting models, the residuals can be obtained by subtracting the observation with the predicted value at time *t*. If the residuals satisfy the assumption of independence, they can be used to construct the control charts.

Step 4: Analyze the residuals

The residuals obtained from the GMDH predicting model must be checked whether the residuals satisfy the assumptions of normality and independence. If the residuals violate the assumptions, go back to Step 3 to modify the GMDH predicting model; if the residuals satisfy the assumptions, go to Step 5.

Step 5: Construct the multivariate control charts

Use the residuals to construct Hotelling's T^2 and MC1 control charts. If one of the Hotelling's T^2 and MC1 control charts reveals that the process is out of control, go to Step 6; otherwise, the Hotelling's T^2 and MC1 control charts show that the process is in control, and the Hotelling's T^2 and MC1 control charts can be adopted for constructing on-line statistical process control limits. Go to Step 1 of phase II.

Step 6: Deal with the out-of-control data points

Three ways to deal with the data points that fall outside the control limits are:

1. If one or two data points fall outside the control limits and the reason of the data points falling outside the control limits can be found and eliminated, then the data points can be eliminated, go back to Step 5 to construct new multivariate control charts.
2. If there are one or two data points falling outside the control limits but the reasons of the data points falling outside the control limit cannot be found, or the reasons can be found but cannot be eliminated, the data points will not be eliminated.
3. If three or more data points fall outside the control limits, it denotes that the process is out of control. The process engineer should find the problems and resolve them. After resolving the problems, go back to Step 1 and recollect the process data.

III. 1. 2 Phase II of the Statistical Process Control Procedure

This subsection presents a procedure to monitor a multivariate autocorrelated process. The procedure is described as follows:

Step 1: Record the information of the process data

Record the process number, machine number, the unit of quality characteristics, the name of operators, and the date of operation, etc. when using the control chart to monitor the process. If the data are recorded in details, the engineers can use this data to analyze the manufacturing process and then correct the process when the process is out of control.

Step 2: Specify the control limits of the control chart

Specify the lower and upper control limits of the Hotelling's T^2 and MC1 control charts constructed in phase I.

Step 3: Sample the data and use the constructed GMDH model to predict the future values of quality characteristics.

Sample the data from each quality characteristic y_{t-N}, ($y_{t-N-1},\cdots,y_{t-1}$) as the process input and use the constructed GMDH model to predict the future value of each quality characteristic, $\hat{y}_t$.

Step 4: Calculate the residuals

Subtract the predicted value obtained from the GMDH predicting model with the actual value for each autocorrelated quality characteristic, the residuals can be obtained.

Step 5: Draw multivariate control charts

Use the residuals to draw the Hotelling's T^2 and MC1 control charts. If there is any data point falling outside the control chart, go to Step 6; otherwise, the control charts denote that the process is in control and go back to Step 1.

Step 6: Diagnose the process and find the problems

If the Hotelling's T^2 control chart reveals that the process is out of control, the MYT decomposition method can be used to find the quality characteristic that causes the process out of control. If the MC1 control chart reveals that the process is out of control, the regression method can be used to find the quality characteristic that causes the process out of control.

Step 7: Remove the assignable causes

According to the result of the diagnosis in Step 6, if the assignable causes can be found and removed, go back to Step 1. If no assignable cause is found, implement the EPC procedure to adjust the process.

III. 2 The Engineering Process Control Procedure

Before employing an EPC procedure to adjust the multiple quality characteristics to their target values, a controller utilizing GMDH must be constructed. This study uses *i* input variables $x_1, x_2 \ldots x_i$ as the process input and *j* quality characteristics $y_1, y_2, \cdots, y_j$ as the process output to construct the GMDH controlling model. The EPC procedure is described as follows:

Step 1: Compute the value of the quality characteristic that is out of control

Subtract the value of the quality characteristic which is out of control from its target value, then the value of the process drifting away from its target, $e_t = y_t - T$, can be obtained.

Step 2: Compute the compensative value of quality characteristic

Compute the compensative value of quality characteristic, $\hat{Y}_{t+1} = T - e_t$.

Step 3: Obtain the predicted values of input variables

The predicted values of input variables ($\hat{x}_{1,t+1}$, $\hat{x}_{2,t+1} \ldots \hat{x}_{i,t+1}$) can be obtained using the compensative value of quality characteristic, $\hat{Y}_{t+1}$, and the GMDH controller.

**Step 4: Choose the appropriate input variable as the

controlling variable

There are two criteria to choose the input variables as the controlling variable for adjusting the process. They are:

1. The range of the adjustment of the input variable is the smaller the better. This study defines the index of evaluating the range of the adjustment of the input variable as the ratio of the volume of the adjustment of the input variable and the standard deviation of the input variable.
2. The influence of the input variables on other quality characteristics is the smaller the better.

The controlling variable not only influence the quality characteristic that is out of control, but also influence other quality characteristics that are in control. Therefore, when choosing an input variable as a controlling variable, the influence of the input variable on other quality characteristics must be considered to avoid that the controlling variable influences other quality characteristics.

Step 5: Use the predicted value of controlling variable as a set-point

Use the predicted value of controlling variable as a set-point to adjust the process and then go back to Step 1 of phase II to proceed to monitor the process.

IV. Conclusion

This study develops an integrated SPC and EPC procedure for a multivariate autocorrelated process. The contributions of the proposed method can be summarized as follows:

1. Because the predicting performance of GMDH is good and the application of GMDH is easier than that of time series and neural network, this study uses GMDH to construct a predicting model to transfer the autocorrelated data into uncorrelated data and construct a controller to adjust the multiple quality characteristics to their target values.
2. Engineers with little knowledge of Statistics can apply the proposed procedure easily. This procedure is very helpful in judging the real process conditions.
3. This study adopts the Hotelling's T^2 and MC1 control charts to detect large shift and small shift in the mean at the same time. Therefore, the proposed procedure has better capability to detect the out-of-control data points.

References

[1] Alwan, L.C. & Roberts, H.V. "Time-series modeling for statistical process control," *Journal of Business and Economic Statistics*, 1988, 6(1), 87–95.

[2] Atienza, O.O., Tang, L.C. & Ang, B.W. "A CUSUM Scheme for Autocorrelated Observations," *Journal of Quality Technology*, 2002, 34(2), 187-199.

[3] Ghare, P.H. & Torgerson, P.E. "The Multicharacteristic Control Chart," *Journal of Industrial Engineering*, 1968, 19, 269-272.

[4] Hotelling, H. "Multivariate quality control," *Techniques of Statistical Analysis*, Eisenhart, Hastay, & Wallis Eds., McGraw-Hill, 1947.

[5] Ivakhnenko, A.G. & Madala, H.R. *Inductive learning algorithms for complex systems modeling*, CRC press, 1994.

[6] Jiang, W., Tsui, K.L. & Woodall, W.H. "A New Monitoring Method : The ARMA Chart," *Technometrics*, 2000, 42(4), 399-410.

[7] Kramer, H.G. & Schmid, W. "EWMA charts for multivariate time series," *Sequential Analysis*, 1997, 16(2), 131–154.

[8] Lowry, C.A. & Woodall, W.H. "A Multivariate Exponentially Weighted Moving Average Control Chart," *Technometrics*, 1992, 34(1), 46-53.

[9] Montgomery, D.C. & Mastrangelo, C.M. "Some statistical process control methods for autocorrelated data," *Journal of Quality Technology*, 1991, 23(3), 179–204.

[10] Montgomery, D.C., Keats, J.B., Runger, G.C. & Messina, W.S. "Integrating Statistical Process Control and Engineering Process Control," *Journal of Quality Technology*, 1994, 26, 79-87.

[11] Montgomery, D.C. *Introduction to Statistical Quality Control, 4th eds*., Wiley, 2001.

[12] Pignatiello, J.J. & Runger, G.C. "Comparisons of Multivariate CUSUM Charts," *Journal of Quality Technology*, 1990, 22(3), 173-186.

[13] Ramírez, J.G. "Monitoring Clean Room Air Using Cuscore Charts," *Quality and Reliability Engineering International*, 1998, 14(5), 281-289.

[14] Timmer, D.H., Pignatiello, J.J. & Longnecker, M.T. "The Development and Evaluation of CUSUM-based Control Charts for an AR(1) Process," *IIE Transactions*, 1998, 30(6), 525-534.

[15] Theodossiou, P.T. "Predicting shifts in the mean of multivariate time series process: an application in predicting business failures," *Journal of the American Statistical Association*, 1993, 88(422), 441–447.

[16] Woodall, W.H. & Ncube, M.M. "Multivariate CUSUM Quality-Control Procedures," *Technometrics*, 1985, 27(3), 285-292.