行政院國家科學委員會專題研究計畫 成果報告

製程品質特性為曲線時之監控方法
研究成果報告(精簡版）

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 94-2118-M-009-006-
執 行 期 間 ：94 年 08 月 01 日至 95 年 10 月 31 日
執 行 單 位 ：國立交通大學統計學研究所

計 畫 主 持 人：洪志真

計畫參與人員：博士班研究生-兼任助理：顏家鈴、村碩慧
碩士班研究生-兼任助理：陳瑩琪、馮耀文、蔡明曄

處 理 方 式 ：本計畫可公開查詢

中 華 民 國 96 年 02 月 02 日

目錄

報告內容

一、前言

Statistical process control (SPC) has been successfully proven useful for quality and productivity improvement in many domains, especially in industries. For most of SPC applications, the quality of a process or product is measured by one or multiple quality characteristics. However, some processes are better characterized by profiles or functions. Kang and Albin (2000) described an example of aspartame (an artificial sweetener), in which the quality is characterized by the amount that dissolves per liter of water at different temperatures. The real example considered in Kang and Albin (2000) is a semiconductor manufacturing application involving the calibration of a mass flow controller, in which the performance of the process is characterized by a linear function. Another example presented in Mahmoud and Woodall (2004) aimed at studying the stability of the calibration curve in the photometric determination of Fe^{3+} with sulfosalicylic acid, in which the relationship between the response variable (the absorbance of the solution) and the explanatory variable (the concentration of the solution) is a linear function. For other examples, see Kim *et al.* (2003) and the papers cited therein.

Some profile monitoring methods have been proposed for this type of processes in the literature. Kang and Albin (2000) and Kim *et al.* (2003) proposed several methods for monitoring the process in which the quality is characterized by a linear profile. Shiau and Weng (2004) proposed a nonlinear curve profile monitoring scheme by nonparametric regression. The models used in the above-mentioned research works assume that there is no subject-to-subject variation; in other words, the randomness of the profile data only comes from the errors such as measurement errors. However, in many practical situations, such as the aspartame example and the vertical density profiles (VDP) of particleboards given in Walker and Wright (2002), there indeed exists subject-to-subject variation.

二、研究目的

The main objective of this study is to propose and study some new profile monitoring schemes for the applications with subject-to-subject variation. Both liner and nonlinear profiles are considered. For nonlinear profiles, both the nonlinear (parametric) regression approach and the nonparametric regression approach are considered. For the nonlinear regression approach, analysts need to choose an appropriate parametric regression model. The nonparametric regression is a more flexible approach since no assumptions are made on the functional form of the profiles except the smoothness of the profiles.

三、文獻探討

Kang and Albin (2000) presented two approaches to monitoring linear profiles. The first approach uses a multivariate T^2 control chart to monitor the profile parameters, slope and intercept, simultaneously. The second approach treats the "residuals" of a sample profile－defined as the deviations from the reference profile at some set values (X-values)－as a rational subgroup and uses a combined EWMA/R (exponentially weighted moving average/range) chart for profile monitoring. Kim *et al.* (2003) presented another approach: for each profile, code the X-values of a profile by centering so that the estimators of the *Y*-intercept and the slope of the regression line are independent; then construct two two-sided EWMA charts to monitor the *Y*-intercept and slope separately and a one-sided EWMA chart to monitor the process variation. They combined these three charts and called this scheme $EWMA_3$.

For nonlinear profiles, Walker and Wright (2002) proposed additive models to assess the sources of variation of vertical density profiles of particleboards. Ding *et al.* (2006) proposed using nonparametric procedures to perform Phase I analysis for multivariate nonlinear profiles. Jin and Shi (2001) used wavelets to monitor and diagnose process faults. Gardner *et al.* (1997) use some spatial signature metrics defined for measuring the deviation of the observed profile from the reference profile to diagnose the equipment faults. They reported that these metric-related charts are very powerful in detecting standard deviation shifts. With a nonlinear regression model, Williams *et al.* (2003) studied the use of Hotelling T^2 control charts to monitor the parameter estimates of the nonlinear regression fits to the successive sets of profile data. Williams *et al.* (2006) gave an application of nonlinear profile monitoring to dose-response data. Jensen and Birch (2006) proposed using a nonlinear mixed model to model profiles in order to account for the correlation structure within a profile.

As to the nonparametric regression approach in profile monitoring, many nonparametric regression estimation methods are available, including, for example, the popular kernel estimation, smoothing splines, local polynomial regression, and spline regression. For nonparametric regression, readers are referred to, for example, the books by Wahba (1990), Hardle (1990, 1991), Hastie and Tibshirani (1990), Green and Silverman (1994), Simonoff (1996), Eubank (1999), and papers cited therein. In this project, for the profile modeling, we adopt spline regression as the curve-fitting/smoothing technique for its simplicity.

四、研究方法

In control-chart implementation for SPC applications, it usually consists of two stages, Phase I and Phase II. A major goal of Phase I is to collect a set of in-control data to establish an in-control baseline for Phase II to construct appropriate control limits for on-line process monitoring. Thus, in Phase I, recognizing out-of-control data from the historical dataset accurately is important. In Phase II, the main concern is to detect process shifts from the baseline established in Phase I as quickly as possible when monitoring on-line data. Thus, to evaluate a monitoring scheme, the main concern in Phase I is to assess how effective the scheme can detect the out-of-control profiles correctly, while the emphasis of Phase II is on detecting process changes as quickly as possible. In Phase I, we use the “true-alarm rate” (the rate of detecting real out-of-control profiles) and the “false-alarm rate” (the rate of claiming in-control profiles out of control) as our assessing criteria. In Phase II, the ARL is often used to compare the performances of competing control charts. Note that the monitoring statistics across the historical profiles are not independent. Thus, we use the Bonferroni method to control the familywise error rate (FWER), which is referred to as the overall false-alarm rate in SPC context. As it is well known, the Bonferroni method can be quite conservative sometimes, which may cause a dramatic loss in detecting power. To meet our goal in Phase I and to enhance the detecting power, we also consider controlling the false discovery rate (FDR) proposed by Benjamini and Hochberg (1995) instead of controlling the FWER. In this study, we extend the procedure provided in Benjamini and Hochberg (1995) to deal with multiple charts in Phase I profile monitoring and call it the Multiple FDR method.

(1) For random-effect linear profile monitoring, we propose combining three Shewhart-type control charts to monitor linear profiles. A combined chart means we monitor the process by more than one chart and the combined chart signals out of control if any of the charts signals. In Phase II operation, since the three statistics under monitoring are mutually independent, we give the same in-control false-alarm rate to each control chart such that the overall in-control false alarm rate of the combined-chart scheme is controlled at the prescribed level α. In Phase I operation, estimators across profiles are not independent, but the three statistics under monitoring for each profile are mutually independent. Therefore, to control the overall false-alarm rate in Phase I monitoring, the Bonferroni method and the Multiple FDR method are implemented and compared. For details, see Chen (2006), a master thesis research work supported by this project.

(2) For nonlinear random-effect profile monitoring with nonlinear regression approach, we propose a new robust method for Phase I operation. In the Phase I analysis of historical data, in order to improve the ability of detecting multiple outliers, we propose using a Hotelling T^2 chart based on the Minimum Covariance Determinant (MCD) estimators proposed by Rousseeuw (1984), which are robust estimator of multivariate location and scale, in conjugation with FDR, which is a relatively new statistical procedure that bounds the number of mistakes made when performing multiple hypothesis tests. For details, see Feng (2006), a master thesis research work supported by this project.

(3) For nonlinear random-effect profile monitoring with nonparametric regression approach, based on Shiau and Weng (2004), we extend the fixed-effect model to a random-effect model in order to provide more variability that we often observe in many profile data, e.g., the aspartame example and VDP example. Like Shiau and Weng (2004), we use the B-spline regression technique to smooth profiles. With the random effect model, we put emphasis on the covariance structure. To analyze the covariance matrix, it is natural to consider the technique of the principal component analysis (PCA). The PCA is very useful in summarizing and interpreting a set of profile data with the same equally spaced values of the

independent variable X for each profile. In this study, we construct our monitoring schemes by utilizing the eigen-vectors and principal component scores obtained from PCA. If these scores correspond to separate modes of variation, then it is natural to monitor each principal component score for what it represents. If not, we suggest a combined-chart scheme that combines the charts of these principal component scores for profile monitoring. In the historical analysis of Phase I data, due to the dependency of principal component scores, we adopt the Hotelling T^2 chart to check for stability. For details, see Tsai (2006), a master thesis research work supported by this project.

Simulation studies are conducted to investigate the effectiveness of the proposed methods and also to compare their performances in terms of the false-alarm rate, detecting power, as well as average run length. Profiles mimicking the aspartame profiles mentioned before are used as an illustrative example.

五、結果與討論

(1) For liner profile monitoring, we propose monitoring statistics for intercept, slope, and variance of the error term. The distributions of these monitoring statistics are derived. The control limits are then established based on these distributions accordingly. The ARL of the combined-chart scheme for Phase II operation is then derived, assuming the process parameters are known. In practice, the in-control values of the process parameters are not known, so a major task in Phase I is to estimate them using historical data collected from the process. To detect out-of-control profiles among the k profiles in the historical dataset, k multiple tests are performed simultaneously. To control an overall false-alarm rate at level α, we must adjust the false-alarm rate for each individual test. Here, we consider two methods, one is the commonly used Bonferroni method and the other is the Multiple FDR method proposed particularly for the combined-chart scheme. Simulation results show that the Multiple FDR method is better than the Bonferroni method in terms of detecting power, especially when more out-of-control profiles are in the historical data. The tradeoff of using the Multiple FDR method is the slightly larger “false-alarm rate”, which is defined as the proportion of the false alarms within the in-control profiles in the historical data.

(2) For nonlinear random-effect profile monitoring with nonlinear regression approach, we propose a profile monitoring scheme which involves first a preprocessing step of the nonlinear regression estimation, then use the parameter estimates to compute a robust T^2 statistic based on MCD estimates, and finally apply the FDR procedure to identify out-of-control samples. Simulation studies show that the proposed FDR scheme (with MCD) outperforms the regular T^2 control chart based on MCD (without FDR). We apply the proposed method to a hypothetical aspartame example to evaluate the performance of the proposed method. Simulation studies show that our methods are effective in detecting any reasonable number of outliers.

(3) For nonlinear random-effect profile monitoring with nonparametric regression

approach, we show by simulation that the sample-covariance-based T^2 chart performs better than the successive-difference-based T^2 chart for temporal shifts. In addition, a simulation study is conducted to evaluate the performance of each principal-component-score chart and the combined-chart in terms of the average run length. The simulation study demonstrates that the combined chart scheme is comparable with the best principal-component-score chart. Also, it is found that the number of principal component scores used in constructing control charts has an effect on the detecting power. We adopt the cross-validation method to choose the number of principal component scores.

The profile monitoring is a useful SPC technique and a promising area of research. More statistical methods, models, and ideas are needed. Curve data analysis techniques given in Ramsay and Silverman (2005) may be useful here.

參考文獻

1. Benjamini, Y. and Hochberg, Y. (1995). "Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing". *J. R. Statist. Soc. B*, 57, 1, 289-300.

2. Chen, Y.-C. (2006). "Monitoring Linear Profiles based on a Random-effect Model". Master Thesis, Institute of Statistics, National Chiao Tung University.

3. de Boor, C. (1978). *A Practical Guide to Splines*. Springer, Berlin.

4. Eubank, R. (1999). *Nonparametric Regression and Spline Smoothing*, 2nd Edition. Marcel Dekker, New York .

5. Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications.* Chapman & Hall, New York.

6. Feng, Y.-W. (2006). "A New Robust Method for Phase I Monitoring of Nonlinear Profiles". Master Thesis, Institute of Statistics, National Chiao Tung University.

7. Gardner, M. M., Lu, J. C., Gyurcsik, R. S., Wortman, J. J., Hornung, B. E., Heinisch, H. H., Rying, E. A., Rao, S., Davis, J. C., and Mozumder, P. K. (1997). "Equipment Fault Detection Using Spatial Signatures". *IEEE Transactions on Components, Packaging, and Manufacturing Technology – Part C*, 20, 295-304.

8. Green, P. J. and Silverman, B. W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach.* Chapman and Hall, New York.

9. Hardle, W. (1990). *Applied Nonparametric Regression.* Cambridge University Press, New York.

10. Hardle, W. (1991). *Smoothing Techniques with Implementation in S.* Springer-Verlag, New York.

11. Hastie, T. and Tibshirani, R. (1990). *Generalized Additive Models.* Chapman and Hall, New York.

12. Jensen, W. A. and Birch, J. B. (2006). "Profile Monitoring via Nonlinear Mixed Models". Technical Report No. 06-4, Department of Statistics, Virginia Polytechnic Institute and State University.

13. Jin, J. and Shi, J. (2001). "Automatic Feature Extraction of Waveform Signals for In-Process Diagnostic Performance Improvement". *Journal of Intelligent Manufacturing*, 12, 257-268.

14. Kang, L. and Albin, S. L. (2000). "On-Line Monitoring When the Process Yields a Linear Profile". *Journal of Quality Technology,* 32, 418-426.

15. Kim, K., Mahmoud, M. A., and Woodall, W. H. (2003). "On The Monitoring of Linear Profiles". *Journal of Quality Technology* 35, pp. 317-328.

16. Mahmoud, M. A. and Woodall, W. H. (2004). "Phase I Analysis of Linear Profiles With Calibration Applications". *Technometrics*, 46, 4, 380-391.

17. Miller, A. (2002). "Analysis of Parameter Design Experiments for Signal-Response Systems". *Journal of Quality Technology* 34, 139-151.

18. Nair, V. N., Taam, W., and Ye, K. Q. (2002). "Analysis of Functional Responses from Robust Design Studies". *Journal of Quality Technology* 34, 355-370.

19. Ramsay, J. and Silverman, B. (2005). *Functional Data Analysis, 2nd Edition.* Springer-Verlag, New York.

20. Rousseeuw, P. J. (1984). "Least Median of Squares Regression". *Jornal of the American Statistical Association*, 79, 871-880.

21. Shiau, J.-J. H. and Weng, Z.-P. (2004). "Profile Monitoring by Nonparametric Regression". Technical Report, Institute of Statistics, National Chiao Tung University. NSC92-2118-M-009-012**.**

22. Simonoff, J. S. (1996). *Smoothing Methods in Statistics*. Springer-Verlag, New York.

23. Tsai, M.-Y. (2006). "Monitoring Random-effect Profiles by Nonparametric Regression". Master Thesis, Institute of Statistics, National Chiao Tung University.

24. Wahba, G. (1990). *Spline Models for Observational Data*. CBMS-NSF Series. SIAM, Philadelphia.

25. Walker, E. and Wright, S. P. (2002). “Comparing Curves Using Additive Models”. *Journal of Quality Technology* 34, pp. 118-129.

26. Williams, J. B., Woodall, W. H., and Birch, J. B. (2003). “Phase I Analysis of Nonlinear Product and Process Quality Profiles". Technical Report No. 03-5, Department of Statistics, Virginia Polytechnic Institute and State University.

27. Williams, J. B., Woodall, W. H., and Ferry, N. M. (2006). “Statistical Monitoring of Heteroscedastic Dose-Response Profiles from High-Throughput Screening”. Technical Report, Department of Statistics, Virginia Polytechnic Institute and State University.

計劃成果自評

本計劃之執行相當順利，結果亦相當豐碩，三位碩士班學生和兩位博士班學生在工業統計上得到相當不錯的訓練。研究內容與原計畫相符程度極高，亦已達成預期目標。研究成果預計將有三篇論文可以在國際知名期刊發表。本研究所提出之剖面資料控制圖在學術上和應用上均有貢獻。