Quality Improvement in the Multi-response Problem by Using Clustering Characteristic

Hui-Mei Yang
Department of Business Administration
Tatung Institute of Commerce and Technology
253, Mi-Tuo Rd. , Chia-Yi, Taiwan 600, R.O.C.

Kun-Lin Hsieh
Department of Information Management
National Taitung University
684, Sec.1, Chung-Hua Rd., Taitung, Taiwan 950, R.O.C.

Lee-Ing Tong
Department of Industrial Engineering and Management
National Chiao Tung University
1001 Ta Hsueh Road, Hsinchu 300, Taiwan R.O.C.

Abstract: - Most previous studies have only addressed a single-response problem. However, more than one correlated response frequently occurs in a manufacturing product. The multi-response problem has received limited attention. In the second part of this project, an approach based on the clustering analysis (CA) is studied to the optimization of the multi-response problem. In CA, the observations can be combined into groups or clusters such that each group or cluster is homogeneous or compact with respect to certain characteristics and each group is different from other groups with respect to the same characteristics. The optimum parameters' settings for a multi-response problem can be determined by three criterions.

Key-Words: - clustering analysis, quality improvement, parameter optimization

1 Introduction

Market demands have impeled manufacturing companies to enhance their product's quality. Off-line quality control is a cost-effective means to optimize the product and process design in support of on-line quality control. A robust design is desired to obtain the optimum design parameter's settings for a product or an operational process in such a manner that the product characteristic or response attains its desired target with minimum variation. Up to now, most related investigations of robust designs are focused on optimizing single response of a manufactured product or process. However, many manufactured products are diversified and this situation causes more than one response to be considered. Furthermore, these responses are usually correlated. For instance, the semiconductor manufacturing or a chemical process must frequently optimize a multiple response problem.

The conventional designed experimental techniques can be employed to study the relationship between the quality response and design parameters (or noise parameters). In addition, Taguchi's method, which combines experimental design techniques with quality loss considerations, is an efficient approach for off-line quality control when the single quality response is involved. To optimize a multi-response problem, multivariate analysis of variance (MANOVA) and the response surface method (RSM) are two methods frequently employed by the analysts. When Taguchi's methods are employed to optimize the multi-response problem, the conflicts are frequently occurred for determining the optimum parameter's settings. Another approach to solve this problem entails assigning a weight for each response. Nevertheless, determining a definite weight for each response in an actual case still remains difficult. However, the possible correlations among the responses may still not be considered. In addition, a factor which has significant effect in a single-response may still not be significant when considered in a multi-response case.

The second part of this project intends to employ the clustering analysis (CA) to perform the optimization of the multi-response problem for both experimental types: the Taguchi's experiments and the conventional experiments. The rest of this project is organized as follows. Section 2 reviews the CA. Section 3 describes the developed optimization procedure. Section 4 provides two numerical examples to demonstrate the effectiveness of the optimization procedure. Conclusions are finally made in Section 5.

2 Literature Review

Derringer and Suich [19] applied the desirability function to optimize the multi- responses problems in a static experiment. Castillo, Montgomery and McCarville [17] demonstrated the modified desirability functions for optimizing the multi-response. However, their method may lead to an inaccurate result for some inexperienced users and may increase the uncertainty in determining the optimal parameter setting, and is difficult for the practitioners who have only limited statistical training. Layne [22] presented a procedure, which considers simultaneously three methods: weighted loss function, desirability function, and a distance function, to determine the optimum parameter combination. The controversies may be generated by simultaneously comparing three methods to determine the optimum setting.

Khuri and Conlon [20] proposed a procedure, based on a polynomial regression model, to simultaneously optimize several responses. Logothetis and Haigh [24] also optimized a five-response process by utilizing the multiple regression technique and the linear programming approach. These two methods are also computationally complex and, therefore, are difficult to be utilized on the shop floor. Pignatiello [25] utilized a variance component and a squared deviation-from-target to form an expected loss function to optimize a multiple response problem. This method is hard to implement for that a cost matrix must be obtained, in addition, the amount of the experimental observations are required. Chapman [18] proposed a co-optimization approach, which composites all response by using a composite response. This approach might confuse some inexperienced practitioners in determining which ranges of the constraint's can be safely expanded.

Leon [23] presented a method, which is based on the notions of a standardized loss function with the specification limits, to optimize a multi-response problem. However, only the nominal-the-best (NTB) characteristic is suitable to employ this approach, which may limit the capability for this approach. Ames et al. [16] presented a quality loss function approach in the response surface models to deal with a multi-response problem. The basic strategy is to describe the response surfaces with experimentally derived polynomials, which can be combined into a single loss function by using known or desired targets. Next, minimizing the loss function with respect to process inputs locates the best operating conditions.

Lai and Chang [21] propose a fuzzy multi-response optimization procedure to search for an appropriate combination or process parameter settings. A strategy of optimizing the most possible response values and minimizing the deviation from the most possible values is used which considers not only the most possible value, but also the imprecision of the predicted responses. Tong et al. [28] developed a multi-response signal to noise (MRSN) ratio, which integrates the quality loss for all responses, to solve the multi-response problem. Conventional Taguchi method can be applied based on MSRN. The optimum factor/level combination can be obtained. Su and Tong [26] also proposed a principle component analysis approach to perform the optimization of the multi-response problem. Initially, standardizing the quality loss of each response; the principle component analysis is then applied to transform the primary quality responses into fewer quality responses. Finally, the optimum parameter combination can be obtained by maximizing the summation standardized quality loss. Tong and Su [27] proposed a procedure, which applied fuzzy set theory to multiple attribute decision making (MADM) for optimizing a multi-responses problem. Although their method can reduce the uncertainty in determining each response's weight, it is still computational complicated to be practically used.

3 Proposed Approach

In this section, we intend to develop a feasible approach to the optimization of the multi-response problem. For a manufactured product or an operational process, the target of the response is usually known. The designed experiments like as the Taguchi's orthogonal array (OA) or the fractional factorial experiments are used to search the optimum parameter's settings. By analyzing the experimental data, an optimum parameter's setting can be determined. For the parameter's settings performed in a designed experiment, it seems to be combined into several groups or clusters. That is, the responses of some groups or clusters are close to the target and that of the others departures from the target. Hence, the cluster analysis can be employed to study the structure of feasible space. Figure 1 depicts graphically the viewpoint.

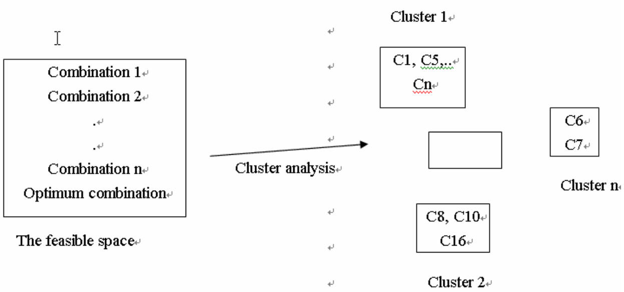

Figure 1.The viewpoint of employing the cluster analysis to the optimization of multi-response problem.

The optimization procedure is described as follows.

Step1.Determine the target response of the multiple responses.

Step2.Determine the significant factors for each response.

Step3.Perform the cluster analysis and determine the number of cluster.

Step4.Determine the optimum parameter's combination by using the following three criterions in the obtained cluster involving target response.

Criterion1: keep the same factor/level to be the optimum level combination.

Criterion2: If there is conflict under decision-making. Significance of factor is initially considered. The optimum factor/level combination of each response must be initially determined. The optimum level can be determined regarding the corresponding factor being significant factor with positive effect on response.

Criterion3: If the optimum level still can not be determined after using criterion 1 and criterion 2, the optimum level can be determined by using the minimum difference between the target responses and the corresponding responses.

4 Propose Approach

Numerical example 1

This numerical example can be referred to Castillo et al [1]. The wire bond heating system has three control factors: flow rate (A), flow temperature (B), block temperature (C). Table 1 lists the levels of each factor. There are six responses to measure three different locations' maximum, beginning, and finishing bonding temperature. The six responses are nominal-the-best(NTB) and the target values are (y1,, y2, y3, y4, y5, y6) = (190, 185, 185, 190, 185, 185). The experimental observations are given in Table 2.

Table 1. The control factors' levels

Factor	Level 1	Level 2	Level 3
A	40	80	120
B	200	325	450
C	150	250	350

Table 2. The experimental observations.

No	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
1	139	103	110	110	113	126
2	140	125	126	117	114	131
3	184	151	133	147	140	147
4	210	176	169	199	169	171
5	182	130	122	134	118	115
6	170	130	122	134	118	115
7	175	151	153	143	146	164
8	180	152	154	152	150	171
9	132	108	103	111	101	101
10	206	143	138	176	141	135
11	183	141	157	131	139	160
12	181	180	184	192	175	190
13	172	135	133	155	138	145
14	190	149	145	161	141	149
15	180	141	139	158	140	148
16	190	185	185	190	185	185

This case is analyzed by the developed optimization procedure. From the experimental observations, there are sixteen data sets (the sixth data set represents the target). These data sets can be viewed as the feasible space. We can search the optimum parameter's settings from the feasible space. Next, the CA is preformed. We use the mixture of the hierarchical and nonhierarchical CA to analyze this data set. From the SAS's output, we consider that the suitable cluster's number to be 4 (the R-square value of four clusters 0.876 is larger than the R-square value of three 0.756). According to the results of SAS, we find that the target, forth and twelve data sets are combined into the same group. It can provide the available information for determining the optimum parameter's settings. The factor/level of both data sets (the forth and the twelve data sets) are (120, 450, 250) and (80, 450, 350). The criterions in step 4 can be then used to determine the optimum settings. For criterion 1, the level of the factor B for both data sets are the same, the level can be kept. According to the information of this case [1], all three factors are significant factors for each response. We should employ the criterion 3 to determine the optimum settings. The level of factor A and B can be determined as 80 and 350 (the response of the twelve experimental run is closer than that of the forth experimental run). Hence, the optimum parameter's settings can be determined as level-type (80, 450, 350). However, the available information will be utilized to determine the continuous-type. From the clustering result, the factor A and factor C should have the continuous optimum settings. For factor A, the continuous optimum settings should lie in the interval (80, 120). For factor C, the continuous optimum settings should lie in the interval (250, 350). Table 3 lists the parameter's settings for Castillo et al and the developed procedure. From this table, we can find that the optimum parameter's settings obtained from CA is very close to the results found from Castillo et al. The parameter's settings found by Castillo et al are the continuous form and the parameter's settings found by CA are discrete form. From the comparisons, they are very close and the possible interval of optimum settings also includes the results of Castillo et al. Hence, the effectiveness and correctness of CA for can be verified.

Table 3. The comparison table.

Approach	Control Factor		
	x1	x2	x3
Castillo *et al*	84.15	450	329.8
Cluster analysis	80	450	350

Numerical 2

This numerical example can be referred to Tong and Su [15]. The PECVD process has eight control factors from A to H. Only factor A has two levels and the other factors have three levels. There are two responses (DT and RI). The two responses are nominal-the-best (NTB) and the target values are (DT, RI) = (1000, 2). The experimental observations are given in Table 4.

This case is analyzed by the developed optimization procedure. From the experimental observations, there are eighteen data sets. We can add one data set, i.e. the target set of the optimum combination. The nineteen data sets can be viewed as the feasible space. Next, we can search the optimum parameter's settings from the feasible space. The CA is preformed. We also use the mixture of the hierarchical and nonhierarchical CA to analyze these data sets. From the SAS's output, we consider that the suitable cluster's number to be 4 (the R-square value of four clusters 0.932 is larger than the R-square value of three clusters 0.845). According to the results of SAS, we find that the target, third and tenth experimental run are combined into the same group. It will provide the available information for determining the optimum parameter's settings. The factor/level of both combinations are A1B1C3D3E3F3G3H3 and A2B1C1D3E3F2G2H1. The criterions of step 4 can be then used to determine the optimum settings. For criterion 1, the level of the factors B, D and E for both combinations are the same, the level's combination (B1D3E3) can then be kept. The optimum factor/level combinations for both responses are: A1B1C3D2E2F2G2H3 to response DT and A1B3C2D1E3F1G1H3 to response RI. According to the information of this case [15], the factors B, C and F are the significant factors for response DT and the factors B, E and F are the significant factors for response RI. The level of factor C can be determined by using the criterion 2 of step 4. The level of factor C can be set to the level 3. Then, we employ the criterion 3 to determine the other optimum settings. The others factors' settings can be determined as A2F2G2H1 by using the criterion 3. After performing the analysis, the final optimum parameter's settings can be represented as A2B1C3D3E3F2G2H1.

Table 4. The experimental observations.

No	DT	RI	No	DT	RI	No	DT	RI
1	730.6	2.033	7	909.8	1.893	13	902.4	1.830
2	874.2	2.224	8	648.8	1.893	14	824.8	2.042
3	967.2	2.611	9	646.6	1.696	15	792.6	2.096
4	800.8	2.018	10	1013.4	1.974	16	814.6	2.192
5	789.2	1.966	11	1293.6	1.828	17	818.0	1.914
6	796.2	1.879	12	900.6	1.896	18	738.8	2.021

Table 3 lists the parameter's settings for Tong and Su, and the developed procedure. From this table, we can find that the optimum parameter's settings obtained from CA still differ with the results found from Tong and Su. The parameter's settings found by Tong and Su considered the priority of responses. Maybe, the priority of response will affect the result. However, it can not be included in this project. We intend to study more detailed in the future.

Table 4. The comparison table (represented as level).

Approach	Control factors							
	A	B	C	D	E	F	G	H
Tong and Su	1	2	3	2	2	2	2	3
Cluster analysis	2	1	3	3	3	2	2	1

6 Concluding Remarks

This study presents an approach to address the quality improvement for a multi-response problem. As for the concept of the proposed, we apply the clustering analysis technique to form the primary core. The useful information about the parameter optimization can be obtained by means of the concept of clustering characteristics. In this study, we also applied two examples to demonstrate the rationality and feasibility of the proposed approach.

References:

[1] Castillo, E. D., Montgomery, D. C. and McCarville, D. R., 1996, "Modified Desirability Functions for Multiple Response Optimization," Journal of Quality Technology, 28(3), p. 337-345.

[2] Logothetis, N. and Haigh, A., 1988, "Characterizing and optimizing multi-response processes by the Taguchi method," Quality and Reliability Engineering International, 4, p. 159-169.

[3] Montgomery, D. C., 1991, Design and Analysis of Experiments, John Wiley & Sons, Inc.

[4] Myers, R. H. and Montgomery, D. C. 1995, Response Surface Methodology, John Wiley & Sons, New York, NY.

[5] Chapman, R. E., 1995-1996, "Photochemistry multiple response co-optimization," Quality Engineering, 8(1), p. 31-45.

[6] Derringer, G. and Suich, R. 1980, "Simultaneous Optimization of Several Response Variables," Journal of Quality Technology, 12(4), p. 214-219.

[7] Johnson, R. A. and Wichern, D. W., 1992, Applied Multivariate Statistical Analysis, Prentice Hall, Inc.

[8] Khuri, A. I. and Conlon, M., 1981, "Simultaneous optimization of multiple responses represented by polynomial regression functions," Technometrics, 23, p. 363-375.

[9] Lai, Y.-J. and Chang, T. H., 1994, "A fuzzy approach for multiresponse optimization: an off-line quality engineering problem," Fuzzy set and system, 63, p. 117-129.

[10] Layne, K. L., 1995, "Methods to determine optimum factor levels for multiple response in designed experimentation," Quality Engineering, 7(4), p. 649-656.

[11] Leon, N. A., 1996-1997, "A pragmatic approach to multiple-response problem using loss function," Quality Engineering, 9(2), p. 213-220.

[12] Phadke, M. S., 1989, Quality Engineering Using Robust Design, Prentice-Hall, Englewood Cliffs, New Jersey.

[13] Pignatiello, J. J., Jr., 1993, "Strategies for robust multi-response quality engineering," IIE Transaction, 25, p. 5-15.

[14] Su, C. T. and Tong, L. I., 1997, "Multi-response robust design by principal component analysis," Total Quality Management, 8(6), p. 409-416.

[15] Tong, L. I. Abd Su, C. T., 1997, "Optimizing multi-response problems in the Taguchi method by fuzzy multiple attribute decision making," Quality and Reliability Engineering International, 13, p. 25-34.

[16] Ames, A. E., Mattucci, N., Macdonald, S., Szonyi, G. and Hawkins, D. M., 1997, "Quality loss functions for optimization across multiple responses surface," Journal of Quality Technology, 29(3), pp. 339-346.

[17] Castillo, E. D., Montgomery, D. C. and McCarville, D. R., 1996, "Modified Desirability Functions for Multiple Response Optimization," Journal of Quality Technology, 28(3), pp. 337-345.

[18] Chapman, R. E., 1995-1996, "Photochemistry multiple response co-optimization," Quality Engineering, 8(1), pp. 31-45.

[19] Derringer, G. and Suich, R. 1980, "Simultaneous Optimization of Several Response Variables," Journal of Quality Technology, 12(4), pp. 214-219.

[20] Khuri, A. I. and Conlon, M., 1981, "Simultaneous optimization of multiple responses represented by polynomial regression functions," Technometrics, 23, pp. 363-375.

[21] Lai, Y.-J. and Chang, T. H., 1994, "A fuzzy approach for multiresponse optimization: an off-line quality engineering problem," Fuzzy set and system, 63, pp. 117-129.

[22] Layne, K. L., 1995, "Methods to determine optimum factor levels for multiple response in designed experimentation," Quality Engineering, 7(4), pp. 649-656.

[23] Leon, N. A., 1996-1997, "A pragmatic approach to multiple-response problem using loss function," Quality Engineering, 9(2), pp. 213-220.

[24] Logothetis, N. and Haigh, A., 1988, "Characterizing and optimizing multi-response processes by the Taguchi method," Quality and Reliability Engineering International, 4, pp. 159-169.

[25] Pignatiello, J. J., Jr., 1993, "Strategies for robust multi-response quality engineering," IIE Transaction, 25, pp. 5-15.

[26] Su, C.-T. and Tong, L.-I., 1997, "Multi-response robust design by principal component analysis," Total Quality Management, 8(6), pp. 409-416.

[27] Tong, L. -I. and Su, C. -T., 1997, "Optimizing multi-response problems in the Taguchi method by fuzzy multiple attribute decision making", Quality and Reliability Engineering International, 13, pp. 25-34.

[28] Tong, L. -I., Su, C.-T. and Wang, C.-H., 1997, "The optimization of multi-response problems in Taguchi method," International Journal of Quality & Reliability Management, 14(4), pp. 367-380.