Detection and Control of Variation Source for a Production Unit

Prof. Jichao Xu[1]
Zhengzhou Institute of Aeronautics
Zhengzhou, Henan, 450005
P.R. China
E-mail:jxu@zzia.edu.cn

Dr. Hasan Akpolat[2]
Faculty of Engineering, City Campus
University of Technology, Sydney
Broadway NSW 2007, Australia
E-mail:hasan.akpolat@uts.edu.au

Abstract

Variation is the archenemy of quality. To reduce or control the variation in a complex production unit, firstly we need to identify the location of the root cause of the variation. This paper discusses the detection of variability and the techniques used for reduction of variation for a production unit consisting of many processes. In the first part of this paper, the background of variability detection in production systems is introduced which is then followed by a weighted network corresponding to correlation matrix of all processes. Based on the network and clustering criterion of maximum spanning tree, a classification of all processes is derived. Furthermore, the variation of each process in a class is determined by residual analysis. In the last part, the use of methods of robust design for the processes with a larger variability is discussed.

Key Words : Variability Source, Clustering, Robust Design, Residual Analysis

1. Introduction

The concept of a product in quality control field means the product on the market, but also includes the part produced under production process. From a traditional quality control point of view, a basic measure criterion for product quality is the ability of a product to meet the tolerance requirements. A product falling within the tolerance limits is classified as a good product and no quality losses will occur. According to this viewpoint, only a product out of the tolerance limits can cause quality losses. However, in recent years this traditional view

[1] A visiting professor in UTS, who is a professor of Zhengzhou Institute of Aeronautics, P. R. China
[2] Senor Lecturer, Engineering Faculty, University of Technology, Sydney (UTS)

has been revised in academic circles as well as industry practices, and a new idea has emerged[1]. According to this new quality viewpoint, the variation of various quality characteristics around the design target values result in quality losses. If one can imagine a situation that there was no variation around the previous design targets in overall production processes, then all products would be perfect. Thus, since 1980s the concept of quality improvement through variation reduction has become an important breakthrough in the ideology of quality engineering. In recent years, academicians have published a huge number of papers about variation reduction, in particular about robustness of a design[2,3]. In 1990s, many world-class companies including Boeing and GM have renewed their quality assurance systems and practices in accordance with the ideas of robust designs.

Usually, it can be assumed that it would be a relatively less complicated task to perform a robust design for a stand-alone process. However in today's modern industry, production systems often are quite complex and consist of many processes with close correlativity. For instance, a piston of automobile requires forty production processes. In the example of the automobile piston, let's imagine that we discovered that the variation of the measurements do not meet the requirements set. Should we conduct the robust design for the last process? The answer to this question is not clear as we do not know which process(es) is/are actually causing the variation. Therefore, a crucial requirement of the robust design in a complex production system is to identify the location of the variation source(s), and then try to eliminate or reduce the variation in that process(es).

This report studies the detection of variability and the corresponding robust design in complex production systems. In section two, the background of variability detection in a production system is introduced. In the following section three, a weighted network corresponding to the correlation matrix of all processes in a production system is set up and a classification of these processes is derived on the basis of the network and clustering criterion of maximum spanning tree. Section four discovers the location of the variability caused by residual analysis. Section six will discuss the methods of robust design that can be effectively conducted for the processes with large variability.

2. The Problems Proposed

For convenience, some concepts are appointed as follows.

Production Unit: The totality of the required processes and their inter-correlation to produce a product is called here a production unit.

The degree of complexity of various production systems will change due to type and nature of products being produced. According to the degree of product complexity, the production systems can be divided into two parts: simple production unit and complex production unit.

Simple Production Unit: A production system that includes either a single process or multiple processes of which are independent each other. See Figure 1.

Figure 1. Single or multiple independent Processes

Complex Production Unit: The production system that consists of multiple processes with correlativity.

In a complex production system, if it has been found that a particular quality characteristic has large variation, due to the correlativity among the processes it is usually difficult to know which process is causing the variation. The correlative relationship between the processes can be categorized into three groups:

(a) Series connection; (b) parallel connection; (c) series-parallel connection (Shown in Fig. 2) Here Y denotes a quality characteristic of the product. In fact, a complex production system may be even more complex than shown here. However, in this paper the authors will concentrate on the above three illustrations as a representation of complex systems.

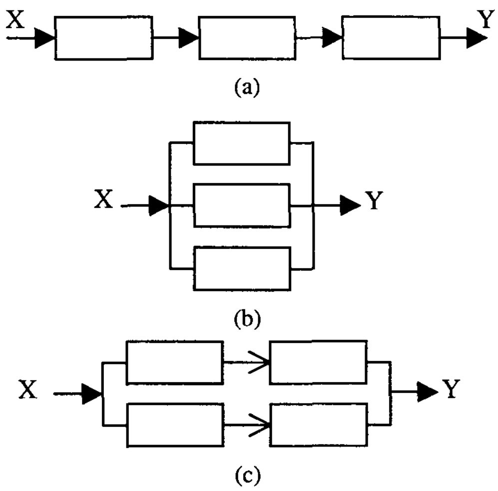


Figure 2

Figure 3 shows a generic model selected for detection of variability in a complex production system.

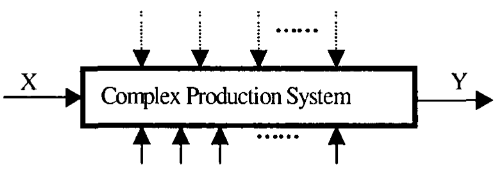


Figure 3. Generic Model for Variation Detection in complex production systems

The random factors, also called noise factors, usually cannot easily be controlled and will affect the variation of product quality characteristic *Y*.

The robust design strategy requires that the settings of system factors, also called control factors, be adjusted so that their performances are insensitive to random factors and as a consequence result in variation reduction. Which system settings or process parameters to adjust is the key question to ask here. Naturally, we need to adjust the system factors in those processes that cause the variation. Therefore, the first task is to locate those processes. For convenience, denote a product quality characteristic or process output by *Y*, and let process quality characteristics in production system should be $Y_1, Y_2 ... Y_n$, respectively. Suppose these quality characteristics are measurable.

3. Study of The Clustering Techniques for Variation Sources

When a process output Y has a large amount of variation, we commonly doubt that all processes contribute in the same way to this variation. However, it is usually difficult to identify which process is actually the source of the variation. The following clustering techniques using determination of residual analysis help screening those processes that may have been the possible source of the variation. For product and each process quality characteristics, *Y* and Y_1, Y_2,... Y_n, the product and process corresponding to *Y* and Y_i are denoted by v_0 and v_1 respectively.

The correlations between product and process, process and process are described by the correlation coefficients of between *Y* and Y_j, Y_i and Y_j $(i \neq j)$, denote the coefficients by ρ_n, ρ_{ij}. $(i, j=1,2,\cdots,n)$. The larger the coefficients are, the closer their correlations. A correlation matrix of product and processes is set up by their correlation coefficients such that

$v_0 \quad v_1 \quad \cdots \quad v_n$

$$\begin{matrix} v_0 \\ v_1 \\ \vdots \\ v_n \end{matrix} \begin{bmatrix} \rho_{00} & \rho_{01} & \cdots & \rho_{0n} \\ \rho_{10} & \rho_{11} & \cdots & \rho_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{n0} & \rho_{n1} & \cdots & \rho_{nn} \end{bmatrix}$$

Equivalently, the correlation matrix can be denoted by a weighted network. In order to cluster, we find out the maximum spanning trees in the weighted network. A valve value, denoted by λ, is determined by the number of clustering. If the number of clustering is large, then λ will also take a large value.

Let the number of classification by clustering technique be k, and each class is denoted by c_i $(i=1,2,\cdots,k)$, then a classification is represented by $\vec{C}=(c_1,c_2,\cdots,c_k)$. Since the solutions of the maximum spanning trees are not unique, the classifications are not unique either. For the class included in the product, $v_0, c_{i0}(i \le i_0 \le k)$, for example, take $c_{i0}=\{v_0,v_2,v_3,v_4\}$.

It has to be mentioned that the correlation among the processes also has to be taken into consideration; perhaps the correlations may be even transmitted to v_0 by a way. Therefore, we use this classification technique to study the relations between and each process. The process quality characteristics corresponding to the branch $\{v_0, v_2, v_3, v_4\}$ are $\{Y, Y_2, Y_3, Y_4\}$. When a product quality characteristic, Y, occurs large variation, it is reasonably viewed that the variation is caused by the larger variation of Y_2,Y_3,Y_4. We can find out their contributions to the overall variation of Y, and determine for which processes robust design should be implemented according to the magnitude of their contributions.

4. The Residual Analysis For Variation Sources

Firstly, we set up the residual analysis model for variability sources. If the variation of Y itself is little after eliminating Y_2,Y_3,Y_4 influence on Y, then it is verified that the processes v_2,v_3,v_4 the variation sources so that a robust design can be conducted for them. For more convenience, let $X=(Y_2,Y_3,Y_4)$, and the residual analysis model be as follows:

$$Y_x = E(Y/x) + \varepsilon_x$$

Where $E(Y/x)$ is the conditional expectation of Y on x, and ε_x is the random error of Y_x

under X=x. A variance $D(\varepsilon_x)=\sigma^2_{\varepsilon_x}$, represents the magnitude of variation from Y itself after eliminating influence Y_2, Y_3, Y_4 on Y. Generally, we assume that ε_x has the same independence distribution for any x, so that the above model can be written as:

$Y = E(Y/x)+\varepsilon$, $D(\varepsilon)=\sigma_\varepsilon^2$.

The linear relationship of $E(Y/x)$ is often assumed (when Y and X has a normal union distribution, and this linearity is true), that is, $E(Y/x)=\beta_0+\beta_1 Y_2+\beta_2 Y_3+\beta_3 Y_4$. Therefore, the residual analysis model is the following:

$$Y=\beta_0+\beta_1 Y_2+\beta_2 Y_3+\beta_3 Y_4+\varepsilon$$

Denoted the variance of Y by $D(Y)=\sigma^2_y$. If $\frac{\sigma^2_y-\sigma^2_\varepsilon}{\sigma^2_y}$ is close to 1, which shows that the variation of Y contributed by Y_2, Y_3, Y_4 is larger, then it is verified that the processes v_2, v_3, v_4 are the variation sources of the product quality characteristic Y. Now, the robust design can be conducted for these processes. In practice, it is not advisable to perform robust design for three or more processes in the same time. Because the variations of Y contributed separately by v_2, v_3, v_4 in the same branch are not equal, we should consider them respectively. Introduce the following model:

$$Y=E(Y/Y_i)+\varepsilon_i \quad \text{i=2,3,4}$$

Calculate, and take the process with the maximum of them as the first object of robust design, and then perform robust design for others in turn. It is not necessary to conduct robust design for the little $\frac{\sigma^2_y-\sigma^2_\varepsilon}{\sigma^2_y}$ because the influence of such Y_i on Y is very subtle. Furthermore, if $\frac{\sigma^2_y-\sigma^2_\varepsilon}{\sigma^2_y}$(i=2,3,4) are close to zero, which it is shown that Y_2, Y_3, Y_4 have little influence on Y, then robust design performed for the processes v_2, v_3, v_4 is not effectively. In this case, remove points v_2, v_3, v_4 in the original network and get a new correlation matrix, weighted network, spanning tree, then make a fresh start to determine a valve value λ and cluster for the new network. According to the above steps to find out new variation sources and perform robust design.

5. Simulation

With the help of Minitabone of the statistical analysis software, we can simulate the overall process of the clustering analysis.
First of all, we assumed that we have such a tandem connection production system:

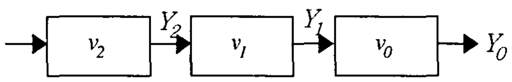


Figure 4

This system has three processes. Figure 4 shows that Y_2 is the input of the process v_1 and the output of the process v_2, Y_1 is the input of the process v_0 and the output of the process v_1 and Y_0 is the output of the process v_0. The following table is data of $Y=(Y_0, Y_1, Y_2)$.

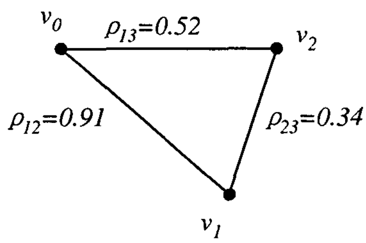
And from this production system, we can get the covariance matrix of Y is as follows:

$$\sum = \begin{pmatrix} 1.0062 & 0.9369 & 0.5127 \\ 0.9369 & 0.9974 & 0.3361 \\ 0.5127 & 0.3361 & 0.9806 \end{pmatrix}.$$

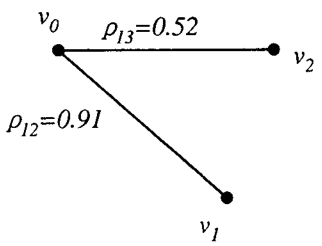
And then we can aquire the correlation matrix of Y.

$$\rho = \begin{pmatrix} 1 & 0.91 & 0.52 \\ 0.91 & 1 & 0.34 \\ 0.52 & 0.34 & 1 \end{pmatrix}$$

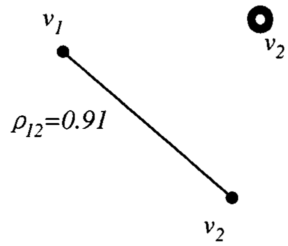
This is a semi-definite matrix. Its weighted network shows in figure 5 (a). In order to cluster, we find out the maximum spanning trees of the weighted network, see figure 5 (b). And then we denote a valve value $\lambda=0.8$



(a) Weighted network



(b) Maximum spanning trees



(c) Clustering

Figure 5

From the above figure 5 (c), we learned that the v_1 and v_2 have a significant relationship. So we assumed that the relationship between v_0 and v_1 is linear. Then the result of the regression analysis is

$$Y_0 = 8.42 + 0.886Y_1 + \varepsilon_x$$

then

$$D(Y_0) = 0.886^2 D(Y_1) + D(\varepsilon_x),$$

$$D(\varepsilon_x) = D(Y_0) - 0.886^2 D(Y_1)$$

where

$$D(Y_0) = 1.006, \quad D(Y_1) = 0.9974,$$

so

$$D(\varepsilon_x) = 1.006 - 0.886^2 \times 0.9974 = 0.19$$

we calculate the $\frac{\sigma^2{}_y - \sigma^2{}_\varepsilon}{\sigma^2{}_y} = \frac{1.006 - 0.19}{1.006} = 0.81$, which is close to 1, so the variation of the Y_0 may be caused by process v_1.

Y_0	Y_1	Y_2
11.1764	2.95763	9.1288
12.4740	5.00382	8.4741
13.5030	5.66514	10.6923
10.9913	2.52535	9.2362
11.2313	3.09364	8.4300
11.7814	4.65450	7.5312
12.5295	4.58385	7.6070
13.4536	5.44226	8.6246
12.0630	4.30360	8.5444
11.9055	3.82046	8.3829
12.1107	4.30810	8.5736
11.7756	3.85018	7.3319
12.5778	4.64678	8.4386
12.0684	3.43097	9.1552
12.2746	3.96462	9.6224
11.6026	3.72472	8.4706
11.3427	2.50003	8.0370
11.5770	3.44297	7.2780
12.9734	5.42369	8.5892
10.3160	2.50149	7.1913
12.7576	4.67417	9.1279
12.6132	4.11655	8.9381
11.5185	3.66243	8.8487
10.6555	2.87993	7.7457
10.8104	2.58749	9.1462
9.9494	2.56124	6.9948
9.9393	1.71094	7.9894
11.4437	3.69982	8.5782
11.4891	3.47643	9.0047
10.7799	3.41584	8.7667
10.3330	2.30978	7.4778
13.4757	4.91223	10.9295
12.4827	3.98084	10.0039
11.8766	3.43593	8.8618
12.2714	4.73221	9.0871
12.7286	4.11670	10.1702

After we have found the source of variation, we can do variation reduction by robust design.

6. The Analysis and Application for Principle of Robust Design

Dr. Taguchi first proposed the concept of robust design and its parameter design technique[3]. It seems that someone often have robust design be equal to parameter design. From the generalized viewpoint, however, robust design can be performed in many ways, of which parameter design is perhaps the most popular method in quality engineering. In this section, we only discuss how to make use of parameter design technique to reduce the variation after the variation source has been detected.

The principle of parameter design of Taguchi, in fact, is a non-linear one. When there exist the determined relation, $Y = f(X)$, between quality characteristics Y and system factor or process parameter X, as shown Fig. 6, the variation of Y can be reduced as the setting of the process parameter X is adjusted to x_1. Therefore the quality of processes or products can be improved by changing the setting of X. When f is of linearity, parameter design will be failure.

In the following, the principle of parameter design can be explained by introducing the viewpoint of conditional random variable. Quality characteristic Y is, in fact, a random variable related to process parameter X. When X is taken different values, Y will be a random variable depending on X. So the conditional random variable sequence or stochastic process, $Y_{x_1}, Y_{x_2}, \cdots, Y_{x_n}, \cdots$ is derived, see Fig 7. Without loss of generality, suppose that Y_X are normally distributed, namely, $Y_x \sim N(\mu_{x_2}, \sigma^2{}_{x_2})$, in which $\sigma^2{}_{x_2}$ is the minimum in all variance of $\{Y_x\}$, so, we should move old setting to new setting x_2. This time the robust design is getting done. From this meaning, the principle of parameter design can be described as: to find a level of process parameter x_0, which can make Y_{x_0} have the minimum variance $\sigma^2{}_{x_0}$. Obviously, when $\{Y_x\}$ are independent and have the same distribution, then parameter design will be failure. The above discussion seems only for univariate parameter, in reality, X can be any vector. This will have the same pattern as the above situation.

There are various implementation techniques regarding parameter design. Taguchi first presented internal and external array methods[3]; we can refer this skill, in which the internal array is used to arrange the combinations of levels of process parameter, namely, the alternatives to be selected, the external array, the employed to calculate the variation of various alternatives. According to the internal and external array, the alternative that is of the minimum variation will be naturally selected. In recent years, many experts and professors in this area create several post-Taguchi methods[2,5], such as simulation analysis method, dual response surface method[2], CAD[2] method etc. When Taguchi method is failure, we can try to use the other methods. Usually, only the variation source is correctly determined, then the parameter design must be effective.

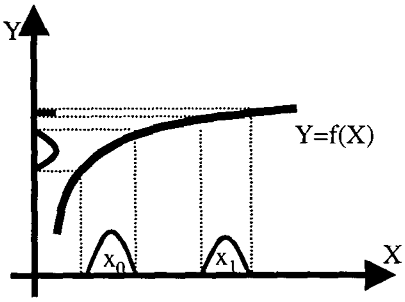


Figure 6

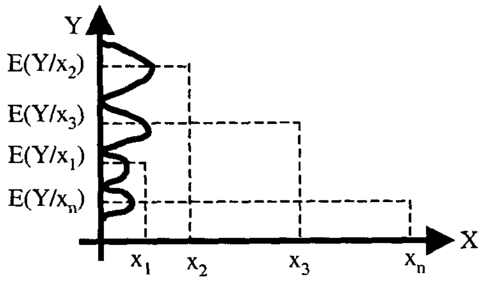


Figure 7

Acknowledgment

This research was jointly supported by National Science Funds for Distinguished Young Scholars 70125004, and NSFC Projects 9870100,70072029 and Engineering Faculty of University of Technology, Sydney. The report was finished during Prof Xu's visit to UTS in 2002. The simulation in this paper are performed by graduate students Dr. Jun Han and Mr Haiyu Wang

References

1. F. Kant Koiper. Defect reduction and process improvement. Quality Breakthrough, a Boeing Commercial Airplane Group Publication for Boeing Suppliers, 1992, Nos. 10 and 11.
2. Vijayan N. Nair. Taguchi's parameter design: A panel discussion, Technometrics, 1992, Vol. 24(2): 127-160.
3. Taguchi G., Systems of experimental design: Engineering methods to optimize quality and minimize cost. White Planins. Ny: UNIPUB/Krans International 1987.
4. Jichao Xu, et.al. The all solutions of clustering of multicriteria. Optimization, Techniques and Applications, Singapore, 1992.
5. Jichao Xu, et.al., Variation Reduction By Data Classification,. The Asian Journal on Quality, 2002, No.2, Vol.2.