A New Approach to Identify Key Quality Characteristics in Multistage Manufacturing Process with Variation Transmission Model

Wang Ning[*]

Business School, Zhengzhou University Zhengzhou China, 450001

Abstract: In this paper, the variation transmission model is introduced to build the key quality characteristics identifying model in multistage manufacturing process. Then the PLSR (Partial Least Squares Regression) method is used to solve the multicollinearity problem of the quality characteristics, make model analyzing and identify the key quality characteristics. At last, the aircraft parts manufacturing process is taken as an example to present the application of this method. The result shows that this method can show the variation transmission relationship of each quality characteristics in every stages and its impaction on the final product quality, and then identify the key quality characteristics, provide the direction for implementing quality control and improvement..

Keywords: Aircraft parts manufacturing, key quality characteristics, multistage manufacturing process, variation transmission.

1. INTRODUCTION

A striking feature of multistage manufacturing process is that the final product quality is determined by complex interactions among multiple stages—the quality variation in a stage are not only influenced by local quality characteristics variation, but also influenced by the variation propagated from upstream stages. In order to identify the key quality characteristics, the most important is to establish the variation transmission model, understand how to increase and delivery variation in multi-stages manufacturing process, and then make quality control and improvement effectively..

In modeling the variation transmission of multi-stage manufacturing process, Lawless, Mackay (1998, 1999) proposed the variation transmission problems earlier. Sui and Otto (1999) proposed a variation transmission model to select the process input parameters which is not sensitive to the process output parameters. Huang *et al.* (2002a) studied the variation identification by applying the state space variation transmission model in multi-stage process. In recent years, Zhou Shiyu, Shi Jianjun, and other scholar had a further research on multistage manufacturing process variation transmission by applying the state space model.

However, the majority scholars fasten on single quality characteristics variation modeling problem in multistage process, but less involves the variation study of multi-quality characteristics. Because in multi-quality characteristics variation analysis, there is not only multicollinearity between quality characteristics, but also a correlation of multi-quality characteristics between each stage, which make the variation formation affected by the multi-correlation, especially in dealing some cases with small sample problem, there is no effective method. So in this paper, the variation transmission model is introduced to build the key quality characteristics identifying model in multistage manufacturing process. Then the PLSR (Partial Least Squares Regression) method is used to solve the multicollinearity problem of the quality characteristics, make model analyzing and identify the key quality characteristics.

2. VARIATION TRANSMISSION MODEL IN MULTI-STAGE MANUFACTURING PROCESS

For the ideal processes, an input must result to a certain output. But, the outputs of a practical process must variety, even though it is thought as steady process, because a practical process must be affected by various outside and inside factors. In fact, the inputs and outputs in any practical process are not same. They must belong to a certain distribution in statistics. These confirm a model about input and output of process (see Fig. **1**).

According to the concept in automatic control, without thought of tools wearing, disturbing outside, adjustment and measurement errors, the function of process input-output model is created as

$$Y = f_s(X) + E \quad \bullet \tag{1}$$

Where $Y = (y_1, y_2, \cdots, y_n)^T$ means the final quality characteristic,

$$X = \left(X_1, X_2, \cdots, X_n \right)^T = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

means the m quality characteristics in n process. $E = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)^T$ is random variation.
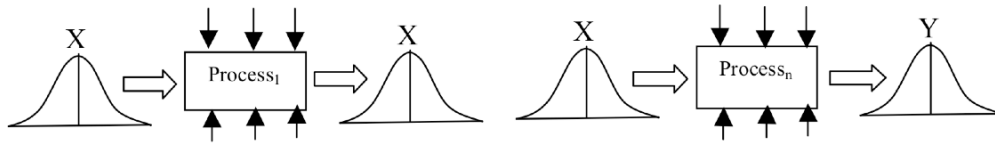
Fig. (1). Variation transmission model of multistage process.

To discuss the input's effect on output of process, the concert of "fore-effect" developed by Zhang is used. The fore-effect output (Y_r) denotes the input's effect on output *via* transmission function. So, the equation (1) could be expressed as:

$$Y = Yr + E \qquad (2)$$

By the formula (1), (2), the variation of process input may be entirely or partly transmitted to process output. The variation of output contains two parts. One part comes from the process itself and another come from the effect of the variation of input. Through the analysis to models, it may be thought that any yi makes of corresponding f(xi) and εi. The distribution of Y is composed of the distributions of "fore-effect" output Yr and variation of process E.

For the confirmed $f_s(\cdot)$ between Y and X, the probability density function, $f_Y(\cdot)$ of Y is confirmed when the $\backslash f_X(\cdot)$ of X is confirmed. The $f_Y(\cdot)$ is expressed as

$$f_Y(y) = f_X[f_s^{-1}(y)]\left|[f_s^{-1}(y)]'\right| \qquad (3)$$

where, $f_X^{-1}(y)$ is inverse function of $f_X(x)$, $[f_X^{-1}(y)]'$ is derivative of $f_X^{-1}(y)$. The PDF of process output can be educed by the convolution..

In practice, usual relation between X and Y is lineal, namely, The above relation between input and output of process can be expressed by lineal function.

$$yr = f_s(x) = \alpha + \beta x \qquad (4)$$

So, the lineal process variation transmission model is

$$Y = \alpha + \beta X + E \qquad (5)$$

where α and β are regression coefficients of lineal model $f_s(\cdot)$. They can be calculated by regression analysis to data of input and output.

The input of process, considered as output of the former process, is usually from normal distribution or approximate. The case of the input data from a normal distribution is discussed in the following, namely $X \sim N(\mu_X, \sigma_X^2)$, where μ_X and σ_X are mean and standard deviation of X respectively.

The above expression shows that the distribution of Yr is normal, $Yr \sim N(\alpha + \beta\mu_X, (\beta\sigma_X)^2)$. So, the relation of mean and variance in case of lineal process variation transmission model can be expressed as

$$\mu_{Yr} = \alpha + \beta\mu_X \qquad (6)$$

$$\sigma_{Yr}^2 = \beta^2\sigma_X^2 \qquad (7)$$

where $\mu_{Yr}, \sigma_{Yr}, \mu_X, \sigma_X$ are means and standard deviation of output Yr and input X respectively.

According to the characters of normal distribution, the convolution of Yr and E, which are normal, is also normal. So, the distribution of output Y is normal in the case of the lineal variation transmission model. Because the variance of output is composed of the variances σ_{Yr}^2 and σ_e^2, the lineal variance transmission model is denoted as ₒ

$$\sigma_Y^2 = \sigma_{Yr}^2 + \sigma_e^2 = \beta^2\sigma_X^2 + \sigma_e^2 \qquad (8)$$

Formula (8) represents the two-stage quality characteristics variation transmission, after recursion, the multi-quality characteristics variation transmission model in multi-stage manufacturing process can be got:

$$\sigma_{ij}^2 = (\beta_{ij}\beta_{i,j-1}\cdots\beta_{i,2})^2\sigma_{1A}^2 + (\beta_{ij}\beta_{i,j-1}\cdots\beta_{i,3})^2\sigma_{2A}^2 + \\ \cdots + \beta_{ij}^2\sigma_{i,j-1,A}^2 + \sigma_{ij,A}^2 \quad i = 1,\cdots,m; j = 1,\cdots,k \qquad (9)$$

where σ_{ij}^2 is the i_{th} quality characteristic variation in the j stage, coefficient matrices β_{ij}^2 represents the transmission degree of the i_{th} quality characteristic variation in the j stage. Here $\sigma_{i1A}^2 = \sigma_1^2$ is the quality characteristic variation in the first stage.

Formula (9) is the key quality characteristics identifying model, if the coefficient matrix $\beta^2 = \begin{bmatrix} \beta_{11}^2 & \beta_{21}^2 & \cdots & \beta_{m1}^2 \\ \beta_{21}^2 & \beta_{22}^2 & \cdots & \beta_{m2}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{1k}^2 & \beta_{2k}^2 & \cdots & \beta_{mk}^2 \end{bmatrix}$ could be estimated, the coefficient matrix can show the variation transmission relationship of each quality characteristics in every stages and its impaction on the final product quality variation, and then identify the key quality characteristics, provide the direction for implementing quality control and improvement.

3. MODELING ANALYSIS AND KEY QUALITY CHARACTERISTICS IDENTIFICATION BASED ON PLSR T

When estimate the coefficient matrix β^2, in order to overcome the multicollinearity between quality characteristics and it with the final product quality, the PLSR method

are introduced to analyze model and identify the key quality characteristics in multistage manufacturing process.

According to previous assumptions, the dependent variable y means the final product quality, independent variable $x_1, \cdots, x_p$ indicate p multistage process quality characteristics, there are n set of independent data samples $(x_{i1}, \cdots, x_{ip}, y_i)$ are collected. After the data pre-treated, it is:

$$y, X = (x_1, x_2, \cdots, x_p)_{n \times p} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}_{n \times p}$$

Firstly, in order to eliminate multicollinearity in X, the integrated variable:

$$t_1 = Xw_1 = w_{11}x_1 + w_{12}x_2 + \cdots + w_{1p}x_p$$

where w_1 is p dimensional column vector, t_1 is n dimensional column vector. t_1 should be have the greatest summarize information of X and get the interpretation ability of Y to achieve maximum, after calculated:

$$w_1 = \frac{1}{\sqrt{\sum_{j=1}^{p} \mathrm{cov}^2(y, x_j)}} \begin{bmatrix} \mathrm{cov}(y, x_1) \\ \vdots \\ \mathrm{cov}(y, x_p) \end{bmatrix}$$

So:

$$t_1 = Xw_1 = \frac{1}{\sqrt{\sum_{j=1}^{p} \mathrm{cov}^2(y, x_j)}} (\mathrm{cov}(y, x_1)x_1 + \cdots + \mathrm{cov}(y, x_p)x_p)$$

Then make regression analysis y and $x_1, \cdots, x_p$ to t_1 respectively: $y = r_1 t_1 + y^{(1)}, x_j = s_{1j} t_1 + x^{(1)}, j = 1, \cdots, p$, where

$$r_1 = \frac{y' t_1}{\|t_1\|}, s_{1j} = \frac{x_j' t_1}{\|t_1\|} \text{ is regression coefficient,}$$

$y^{(1)}, x_j^{(1)}$ is the residual vector in the first regression. Then $y^{(1)}, X^{(1)} = (x_1^{(1)}, \cdots, x_p^{(1)})$ is used to replace the y and X in above formula with the above method and extract the integrated variables again, and also make the regression analysis of $y^{(1)}, X^{(1)}$ to t_2 until extract A integrated variable. A is always determined by "cross-validation" and do multiple regression analysis of $t_1, \cdots, t_A$ to y, was: $\hat{y} = b_1 t_1 + b_2 t_2 + \cdots + b_A t_A$. Due to $t_j (j = 1, \cdots, A)$ is the linear combinations of $x_1, \cdots, x_p$, so can get the regression model: $\hat{y} = (\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p)$.

Then we can substitute the coefficient matrix β Obtained by upper PLS regression into the formula (7), so the

coefficient matrix β^2 can be obtained, sequentially the multi-quality characteristics variation transmission in multistage manufacturing process can be built.

4. CASE STUDY

In this paper, a mandrel parts manufacturing process of aircraft air conditioning equipment in Yubei factory of XH Group is taken as an example to verify the correctness and effectiveness of the identifying methods. There are location, cuttings and drilling, three stages in mandrel production. Due to location and cuttings are two parallel processes, here only location, cuttings and drilling two manufacturing process researched.

According to the stipulation of aircraft production standards in the Group, the mandrel borehole flatness (Y$_1$), the mandrel borehole diameter (Y$_2$), the mandrel borehole roundness (Y$_3$), the mandrel borehole deepness (Y$_4$) are the important indicators of the part inspection quality. So the four indicators are selected as key quality characteristics in the second stage, respectively as the dependent variable. And in location and cuttings, based on historical experience, we select the fixture strength (X$_1$), fixture diameter (X$_2$), slip angle (X$_3$), mandrel diameter (X$_4$), cutter diameter (X$_5$) and cutter stamping strength (X$_6$), cutting tool angle (X$_7$), cutting tool strength(X$_8$) eight quality characteristics respectively as mandrel drilling manufacturing process independent variables, as the first stage quality characteristics. There are 13 batches production data were collected and first, according to the above raw data logarithmic transformation table, a correlativity matrix between all variables is obtained, as shown in Table **1**.

It can be seen from the correlativity matrix that the correlation coefficients between variables are all very high, there is a strong correlation among the variables. Therefore, PLS regression analysis methods need to be used.

In many cases, partial least squares regression equation does not need to use all of the components in regression modeling, but use the truncated way to choose the first m components to build the model. In fact, if the follow-up component cannot provide more meaningful information in explaining the response variables, using more components only will damage the understanding of statistical trends, and guide the error conclusions. Now Cross Validation (CV) method is widely used at home and abroad to determine the key characteristics.

Table **1** Biologically treated effluents (mg/l)[a]

According to the output results of Fig. (**2**), the explanatory power to y$_1$ of extracting one PLS component is 0.061, the second component is -0.0033, the cross validation value of the second component is less than the critical value, as $Q_3^2 < 0.05$, therefore, y$_1$ only extract one component, the explanatory power of model for y$_1$ is 0.311. Similarly, y2, y3, y4 extract 2,1,1 components from the model, their explanatory power, respectively ,are 0.587,0.483 and 0.697, having reached a high accuracy.

The multivariate PLSR model could be got eventually using SIMCA-P software, as follows:

Table 1.　　**Variable correlation coefficient table.**

	Y_1	Y_2	Y_3	Y_4	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Y_1	1											
Y_2	0.196	1										
Y_3	0.125	-0.542	1									
Y_4	0.461	0.628*	-0.166	1								
X_1	0.581*	0.57*	-0.233	0.676*	1							
X_2	-0.606*	-0.305	-0.088	-0.636*	-0.833**	1						
X_3	0.172	-0.14	0.306	0.189	-0.005	-0.339	1					
X_4	-0.19	0.106	-0.167	0.157	0.1	0.003	0.172	1				
X_5	-0.325	-0.728**	0.394	-0.658*	-0.597*	0.4	-0.171	0.515	1			
X_6	0.46	0.515	-0.127	0.804**	0.671*	-0.722**	0.082	0.141	-.637*	1		
X_7	0.259	0.426	-0.293	0.53	0.325	-0.28	0.47	-0.113	-0.49	.366	1	
X_8	0.273	-0.151	0.559*	0.183	0.253	-0.618*	0.344	-0.196	0.072	.354	-.236	1

*. at the 0.05 level (bilateral) is significantly relevance,
*. at 0.1 level (bilateral) is significantly relevance.

Var ID (Primary)		R2VY	R2VY(cum)	Q2VY	Q2 limit	Q2VY(cum)
Total	Comp 1	0.368963	0.368963	0.0445713	0.05	0.0445713
	Comp 2	0.161351	0.530314	0.119809	0.05	0.15904
	Comp 3	0.0302586	0.560573	-0.0115727	0.05	0.149308
Y1	Comp 1	0.310975	0.310975	0.0610638	0.05	0.0610638
	Comp 2	0.0417387	0.352713	-0.00335928	0.05	0.0579097
	Comp 3	0.0738544	0.426568	0.0256714	0.05	0.0820944
Y2	Comp 1	0.413894	0.413894	0.216764	0.05	0.216764
	Comp 2	0.173335	0.587229	0.0700687	0.05	0.271645
	Comp 3	0.0325591	0.619788	-0.0413468	0.05	0.241529
Y3	Comp 1	0.0540404	0.0540404	-0.643253	0.05	-0.1
	Comp 2	0.428932	0.482972	0.291196	0.05	0.220316
	Comp 3	0.0146188	0.497591	-0.00847219	0.05	0.21371
Y4	Comp 1	0.696941	0.696941	0.54371	0.05	0.54371
	Comp 2	0.00140009	0.698341	-0.038928	0.05	0.525948
	Comp 3	2.16564e-006	0.698344	-0.0560624	0.05	0.499371

Fig. (2). PLSR calculations results.

$$y_1 = 35.456 + 0.228x_1 - 0.203x_2 - 0.027x_3 - 0.234x_4 - 0.097x_5 + 0.129x_6 + 0.063x_7 + 0.087x_4$$

$$y_2 = 182.272 + 0.262x_1 - 0.055x_2 - 0.215x_3 - 0.076x_4 - 0.31x_5 + 0.161x_6 + 0.196x_7 - 0.242x_8$$

$$y_3 = 1397.7 - 0.146x_1 - 0.102x_2 + 0.285x_3 + 0.007x_4 + 0.27x_5 - 0.055x_6 - 0.162x_7 + 0.372x_8$$

$$y_4 = 43.767 + 0.232x_1 - 0.173x_2 + 0.002x_3 + 0.027x_4 - 0.238x_5 + 0.225x_6 + 0.169x_7 - 0.007x_8$$

Because this case only considers two-stage condition, so we take the regression coefficient values into formula (7), and can be obtained:

$$\sigma_{y_1}^2 = 0.052\sigma_{x_1}^2 + 0.0412\sigma_{x_2}^2 + 0.00073\sigma_{x_3}^2 + 0.0548\sigma_{x_4}^2 + 0.0094\sigma_{x_5}^2 + 0.0166\sigma_{x_6}^2 + 0.004\sigma_{x_7}^2 + 0.0076\sigma_{x_8}^2$$

$$\sigma_{y_2}^2 = 0.0686\sigma_{x_1}^2 + 0.003\sigma_{x_2}^2 + 0.0462\sigma_{x_3}^2 + 0.0058\sigma_{x_4}^2 + 0.0961\sigma_{x_5}^2 + 0.026\sigma_{x_6}^2 + 0.0384\sigma_{x_7}^2 + 0.0586\sigma_{x_8}^2$$

$$\sigma_{y_3}^2 = 0.0213\sigma_{x_1}^2 + 0.01\sigma_{x_2}^2 + 0.0812\sigma_{x_3}^2 + 0.000049\sigma_{x_4}^2 + 0.0729\sigma_{x_5}^2 + 0.003\sigma_{x_6}^2 + 0.0262\sigma_{x_7}^2 + 0.1384\sigma_{x_8}^2$$

$$\sigma_{y_4}^2 = 0.0538\sigma_{x_1}^2 + 0.03\sigma_{x_2}^2 + 0.000004\sigma_{x_3}^2 + 0.00073\sigma_{x_4}^2 + 0.0566\sigma_{x_5}^2 + 0.0506\sigma_{x_6}^2 + 0.0286\sigma_{x_7}^2 + 0.000049\sigma_{x_8}^2$$

This model is the multi-quality characteristics variation transmission model in cigarette manufacturing process. For y_1, the variation coefficient of x_1, x_2, x_4 are bigger than other quality characteristics obviously, which are the key quality characteristics for mandrel borehole flatness. Similarly available, x_1, x_5, x_8 are the key quality characteristics for the mandrel borehole diameter, x_3, x_5, x_8 are the key quality characteristics for the mandrel borehole roundness, x_1, x_5, x_6 are the key quality characteristics for the mandrel borehole deepness.

CONCLUSION

In this paper, variation transmission model is used to provide a new method in identifying key quality characteristics in multistage manufacturing process. However, this method is still inadequate, firstly paper assumes that measurement error is zero, but measurement error is widespread, in much actual manufacturing process. When measurement error variation is big, it will have an impact on the quality characteristics variation of final product, which is needed to further consideration; Second, multi-stage manufacturing process cannot always be linear representation, which may be more non-linear relationship in common. However, the study on non-linear relationship is still quite rare, future efforts should be made on it,

CONFLICT OF INTEREST

The author confirms that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Daniel Y. T. Fong and J. F. Lawless. "The analysis of process variation transmission with multivariate measurements". Statistica Sinica, pp.151-164,1998.

[2]   J. F. Lawless, R. J. Mackay, J.A. Robinson. "Analysis of Variation Transmission in Manufacturing Process—Part1". Journal of Quality Technology, pp.131-142 vol.4, no.31, 1999.

[3]   R. Agrawal, J. F. Lawless, R. J. Mackay. "Analysis of Variation Transmission in Manufacturing Process—Part2". Journal of Quality Technology, pp.143-154, vol.4, no.31, 1999.

[4]   Suri, R.and K. Otto. "System-level robustness through integrated modeling". ASME International Conference on Design Theory and Methodology. Las Vegas, Nevada, 1999.

[5]   Huang, Q., Zhou, S. and Shi, J. "Diagnosis of multi-operational machining processesthrough variation propagation analysis", Robotics and CIM Journal, Vol. 18, pp.233-239. 2002a.

[6]   li Zeng and Shiyu Zhou. "Variability monitoring of multistage manufacturing processes using regression adjustment methods". IIE Transactions, pp.109-121, no.40, 2008.

[7]   Jean-Philippe Loose, Shiyu Zhou and Dariusz Ceglarek. "Variation Source Identification in Manufacturing Processes Based on Relational Measurements of Key Product Characteristics". Journal of Manufacturing Science and Engineering, Vol. 130,2008.

[8]   Shi, J., Zhou, S. "Quality control and improvement for multistage systems: A survey ". IIE Transactions, pp.744-753, no.41, 2009.

[9]   G.X.Zhang. "A New Type of Control Chart and a Theory of Diagnosis with Control Charts", World Quality Congress Transaction, pp.175-185, 1984.

[10]   JA Heredia, M Gras. "Statistical estimation of variation transmission model in a manufacturing process". The International Journal of Advanced Manufacturing Technology, pp.789-795, vol.52, no. 5-8, 2011.

[11]   Richard O. Montes. "Variation Transmission Model for Setting Acceptance Criteria in a Multi-staged Pharmaceutical Manufacturing Process", AAPS PharmSciTech, pp.193-201, vol.13, no.1, 2012

[12]   Vahid Abedini, Mohsen Shakeri, Mohammad Hasan Siahmargouie. "Error Analysis in Multistage Machining Process Using Kinematic Analysis of Workpiece Fixturing". Applied Mechanics and Materials, pp. 430-435 ,Volumes 152-154,2012