Journal of Industrial and Systems Engineering
Vol. 1, No. 3, pp 251-259
Fall 2007

Applying the Mahalanobis-Taguchi System to Vehicle Ride

Elizabeth A. Cudney[1], Kioumars Paryani[2*], and Kenneth M. Ragsdell[3]

[1] University of Missouri – Rolla, Rolla, Missouri 65409 USA
(elizabeth.cudney@umr.edu)

[2] Lawrence Technological University, Southfield, Massachusetts 02139 USA
(marskiou@yahoo.com)

[3] University of Missouri – Rolla, Rolla, Missouri 65409 USA
(ragsdell@umr.edu)

ABSTRACT

The Mahalanobis Taguchi System is a diagnosis and forecasting method for multivariate data. Mahalanobis distance is a measure based on correlations between the variables and different patterns that can be identified and analyzed with respect to a base or reference group. The Mahalanobis Taguchi System is of interest because of its reported accuracy in forecasting small, correlated data sets. This is the type of data that is encountered with consumer vehicle ratings. MTS enables a reduction in dimensionality and the ability to develop a scale based on MD values. MTS identifies a set of useful variables from the complete data set with equivalent correlation and considerably less time and data. This paper presents the application of the Mahalanobis-Taguchi System and its application to identify a reduced set of useful variables in multidimensional systems.

Keywords: Mahalanobis-Taguchi system (MTS), Mahalanobis distance (MD), Mahalanobis space, Pattern recognition, Signal-to-noise ratio, Orthogonal array

1. INTRODUCTION

The primary objective of this research is to develop a methodology, which demonstrates the relationship between actions of a producer and its suppliers (for instance, target setting of sub-systems or components' performance attributes) and consumer satisfaction ratings. The goal is to efficiently forecast consumer satisfaction ratings (CSRs) as a function of available vehicle level performance data for vehicle ride. The purpose of this research is to develop a relationship between vehicle attributes and measured customer satisfaction ratings (CSRs) for the purpose of understanding and improving customer driven quality. MTS enables a reduction in dimensionality and the ability to develop a scale based on MD values (Taguch and Jugulum, 2002).

Consumer satisfaction ratings are calculated by market research based on consumers' responses to

* Corresponding Author

survey statements. The vehicle ride data is acquired through laboratory testing, design elements, and reverse engineering. In this research, 67 vehicle data sets are considered for the 6 vehicle ride parameters. MTS is employed to develop a relationship between the 6 vehicle ride parameters and their respective consumer satisfaction ratings (CSRs) for the 67 vehicles. Consumers judge quality and performance at the vehicle level, but important cost-effective decisions at the sub-system or component level must be made by the producer in order to economically satisfy consumer's needs by providing affordable and high quality products. Consumers evaluate various vehicle level attributes such as ride, handling, acceleration, braking and roominess. These attributes are influenced by many factors at all levels of the vehicle architecture, and these factors are often correlated.

The design problem would be significantly simplified if vehicle attributes depended on a unique set of vehicle factors (variables). That is, if one could isolate a set of factors for ride, another for handling, and yet others for, acceleration, roominess, etc; then the design of high quality, low cost vehicles would be much easier. Unfortunately, this is not the case. The real situation is much more complex. Factors at the subsystem and part levels often affect several vehicle attributes, and for a single vehicle attribute the controlling factors are often correlated.

An effective methodology is needed to handle this more realistic and complex situation. The proposed methodology uses a pattern recognition scheme known as the Mahalanobis-Taguchi System (MTS) to translate the performance of lower level elements into an estimate of consumer satisfaction rating at the vehicle level (Taguch and Jugulum, 2002).

2. REVIEW OF RELEVANT LITERATURE

Considerable research is available using Mahalanobis distance to determine similarities of values from known and unknown samples. Existing research also uses the Mahalanobis-Taguchi System for prediction and diagnosis which illustrates the methodology's accuracy and effectiveness.

Jugulum and Monplaisir (2002) performed preliminary comparison between MTS and Neural networks. They showed that in the case of large samples both methods perform equally well and in the case of small samples, MTS is somewhat better than neural networks. In this research, they have not compared these methods in terms of reducing the attributes.

A comparison of the Mahalanobis-Taguchi System and Neural Network was also provided in the work of Hong et al. (2005). This research utilized a breast cancer study to compare the ability of MTS and neural network algorithm with varying numbers of attributes and different numbers of data size. The results indicated MTS performed better with small sample sizes than neural network.

Pattern recognition using Mahalanobis distance was demonstrated in the work of Wu (2004). In this research, pattern recognition was used to diagnose human health. The research considered diagnosis of liver function with the objective to forecast serious disease until the next check-up. The approach provided a more efficient method that also avoided inhuman treatment that had previously been used double blind tests.

Asada (2001) used the Mahalanobis-Taguchi System to forecast the yield of wafers. Mahalanobis distances were calculated on various wafers to compare the relationship between yield and distance. The signal-to-noise ratios were used to indicate the capability of forecasting and the effect of the parameters. This research showed the applicability of the Mahalanobis distance to predict the defective components.

3. MAHALANOBIS DISTANCE

Prasanta Chandra Mahalanobis introduced Mahalanobis distance (MD) in 1936 for the first time. The Mahalanobis-Taguchi System (MTS) was later developed by Genichi Taguchi as a diagnosis and forecasting method using multivariate data for robust engineering.

Mahalanobis distance (MD) is a distance measure that is based on correlations between variables and the different patterns that can be identified and analyzed with respect to a reference population. MD is a discriminant analysis tool (Taguch and Jugulum, 2002), which will be used to predict changes in consumer satisfaction corresponding to changes in multiple engineering characteristics at all levels of a hardware set (vehicle design).

MD is a measure based on correlations between variables and the different patterns that can be identified and analyzed with respect to a reference point. MD is very useful in determining the similarity of a set of values from an unknown by comparing a sample from the unknown group to a measured collection of known samples (Lande, 2004).

Traditionally, the MD methodology has been used to classify observations into different groups. MD is defined in Equation (1).

$$MD_j = D_j^2 = \frac{1}{k} Z_{ij}^T A^{-1} Z_{ij} \tag{1}$$

where,

- k = total number of variables
- i = number of variables (i = 1, 2, . . ., k)
- j = number of samples (j = 1, 2, . . ., n)
- Z_{ij} = standardized vector of normalized characteristics of x_{ij}
- Z_{ij} = $(x_{ij} - m_i)/s_i$
- x_{ij} = value of the ith characteristic in the jth observation
- m_i = mean of the ith characteristic
- s_i = standard deviation of the ith characteristic
- ^{T} = transpose of the vector
- A^{-1} = inverse of the correlation matrix

Mahalanobis distance is used to determine the similarity of a known set of values to that of an unknown set of values. MD has successfully been applied to a broad range of cases mainly because it is very sensitive to inter-variable changes in data. Also, since the Mahalanobis distance is measured in terms of standard deviations from the mean of the samples, it provides a statistical measure of how well an unknown sample matches a known sample set (Manley, 1994).

MD is a discriminant analysis tool that is used in this research to translate the lower level functions into an estimate of consumer satisfaction at the product level. MD is used to predict how changes in engineering characteristics in the product design impact consumer satisfaction.

4. THE MAHALANOBIS-TAGUCHI SYSTEM

The Mahalanobis Taguchi System (MTS) is a pattern recognition technology that aids in quantitative decisions by constructing a multivariate measurement scale using a data analytic method. The main objective of MTS is to make accurate predictions in multidimensional systems by constructing a measurement scale (Taguch and Jugulum, 2002). The patterns of observations in a multidimensional system highly depend on the correlation structure of the variables in the system. One can make a wrong decision about the patterns if each variable is analyzed separately without considering the correlation structure. To construct a multidimensional measurement scale, it is important to have a distance measure. The distance measure is based on the correlation between a variable and the different patterns that could be identified and analyzed with respect to a base or reference point.

In the MTS, the Mahalanobis space (reference group) is obtained using the standardized variables of healthy or normal data. The Mahalanobis space can be used to discriminate between normal and abnormal objects. Once this MS is established, the number of attributes is reduced using the tools of orthogonal array (OA) and signal-to-noise ratio (SN) by evaluating the contribution of each attribute. Each row of the OA determines a subset of the original system by the including and excluding that attribute of system.

5. VEHICLE RIDE

The MTS can be used to minimize the number of variables required to forecast the performance of a system. The objective of the MTS analysis for vehicle ride is to explain the relationship between vehicle ride parameters and customer satisfaction rating as reported by a survey. The purpose is to provide an understanding of the relationship between vehicle ride and customer satisfaction. This relationship will offer an opportunity to efficiently improve customer driven quality.

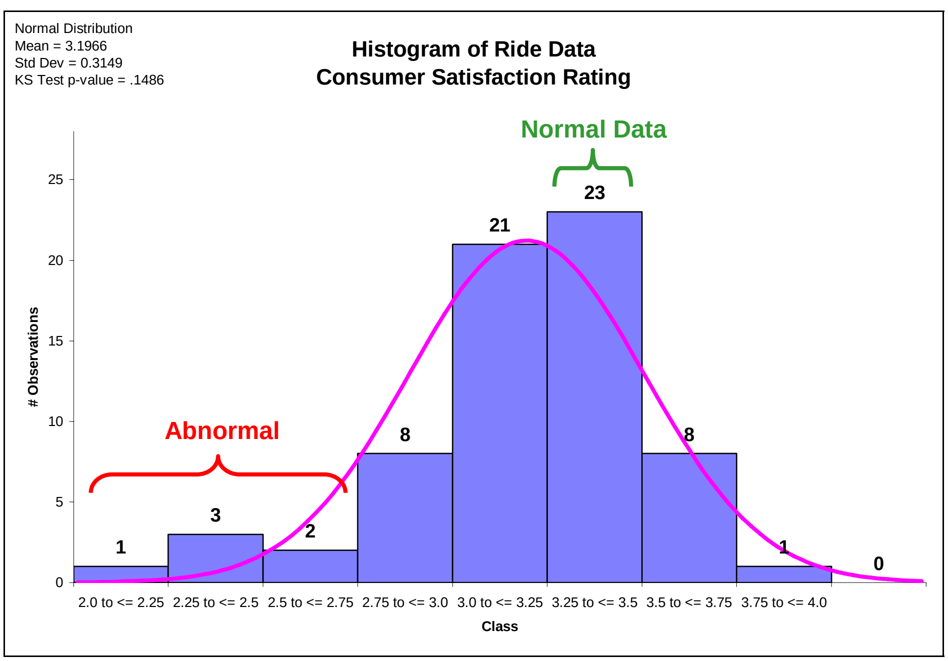

Figure 1. Histogram of Ride Data

The ride data consisted of 67 vehicle data sets and six factors. The first step in performing MTS analysis is to define a “normal” or “healthy” group as the Mahalanobis space. The normal group is selected with discretion to define a reference point on a measurement scale. Defining the normal group is a critical step in this method since the MS is the reference point and base of the measurement scale. An abnormal condition is outside of the healthy group. The degree of abnormality is measured in reference to the normal group.

Table 1. Normal Group MD Values

Data Set	**CSR**	**MD**
26	3.28	0.987
31	3.28	0.612
53	3.28	0.605
1	3.30	1.857
67	3.31	1.549
17	3.33	1.035
23	3.33	0.871
10	3.34	1.169
18	3.35	0.679
62	3.35	2.067
64	3.36	0.682
7	3.37	0.366
52	3.37	0.428
34	3.38	0.407
60	3.38	1.693
21	3.39	0.706
63	3.41	0.931
48	3.42	0.814
35	3.43	1.431
59	3.45	1.690
5	3.47	0.934
46	3.47	0.982
15	3.48	0.505

A summary of the MD values for the normal group and test group is provided in Figure 2.

Table 2. Test Group MD Values

Data Set	**CSR**	**MD**
25	2.16	10.858
37	2.40	6.074
47	2.46	8.309
16	2.48	4.348
55	2.68	2.545
2	3.30	8.001

A histogram using the measured consumer satisfaction rating for the 67 data sets was constructed to select the normal group. Twenty-three samples were selected for the normal group. The samples were selected as the largest grouping of samples. Six samples that fell outside of the normal group were selected as the test group. Figure 1 shows a histogram of the 67 data sets.

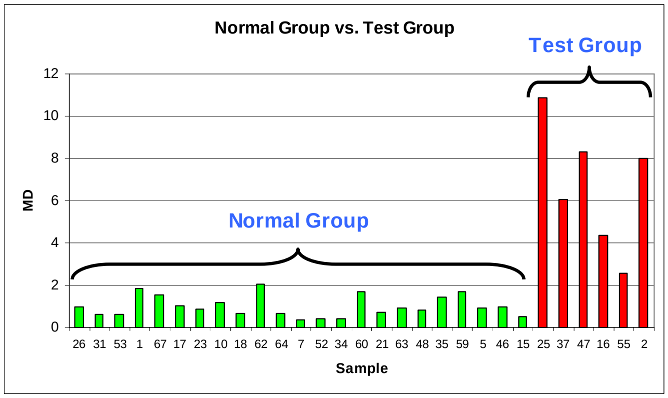

Figure 2. MDs for the Normal Group and Test Group

The Mahalanobis distance was then calculated using the data for the normal group. Table 1 shows the resulting MD value for each sample from the normal group. The data is listed in order of ascending consumer satisfaction ratings rating. The MDs were then calculated for the test group. Table 2 summarizes the MDs. The data sets are listed in order of ascending CSR.

The next step is to optimize the discriminating ability of the MTS system using orthogonal arrays and signal-to-noise ratio. In this step, the purpose is to reduce the dimensionality of a multivariate system but still obtain meaningful results. Orthogonal arrays (OA) are used to estimate the effect of factors and interactions by minimizing the number of experiments required (Kiemele et al, 1999). The appropriate 2 level OA is selected based on the number of factors and number of levels (Fowlkes and Creveling, 1995). An OA is employed in the MTS to minimize the number of runs in an experiment. Using an OA also allows each factor an equal opportunity of impacting the system. Level 1 in the OA means the factor should be included to construct the Mahalanobis space. Level 2 means the factor should not be included to construct the Mahalanobis space. Table 3 shows the L8 orthogonal array used for the vehicle ride analysis. For example, in the second run (1, 1, 1, 2, 2, 2) only the factors X_1, X_2, and X_3 are used to construct the space. Factors X_4, X_5, and X_6 are not included in the MD. Only three factors are considered in the calculation.

Table 3. L_8 Orthogonal Array

	X_1	X_2	X_3	X_4	X_5	X_6
1	1	1	1	1	1	1
2	1	1	1	2	2	2
3	1	2	2	1	1	2
4	1	2	2	2	2	1
5	2	1	2	1	2	1
6	2	1	2	2	1	2
7	2	2	1	1	2	2
8	2	2	1	2	1	1

For each experiment in the OA a signal-to-noise (S/N) ratio is calculated. S/N ratios measure the effect of including or excluding a factor in the MTS. The S/N ratio is an indication of the propagation of variation. The S/N ratio, obtained from the test MDs, is used as the response for each combination of OA. Based on the signal-to-noise ratios calculated using the runs in the orthogonal array, analysis of means (ANOM) tables are constructed to determine the useful variables and identify the candidate variables for elimination. For S/N ratios, the larger the number in the positive direction indicates a greater impact of the corresponding variable on the system.

The useful set of variables is determined using orthogonal arrays and the signal-to-noise ratio. The useful set of variables is obtained by evaluating the gain in the S/N ratio. For the vehicle ride, five of the six parameters are turned out to be useful variables. The gains for the ride parameters are shown in Figure 3.

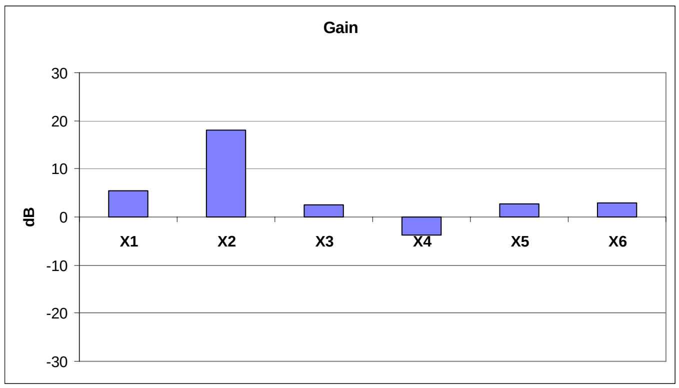

Figure 3. Ride Gains

Regression analysis is then performed. Using only the top two useful variables, a correlation coefficient of 0.864 is achieved. Using the regression equation, the predicted consumer satisfaction rating values are calculated. The predicted values are compared to the actual consumer satisfaction rating values as graphically shown in Figure 4.

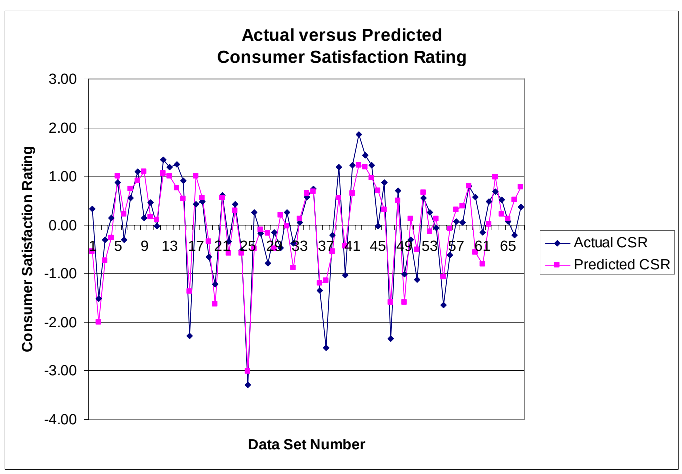

Figure 4. Measured Versus Predicted CSR for Top Two Useful Variables

Scatter plots were then developed to show the correlation between the measured versus predicted consumer satisfaction rating values as shown in Figure 5. The scatter plot is useful to show the linearity of the prediction model. As useful variables are added to the prediction model, the slope and prediction accuracy increase.

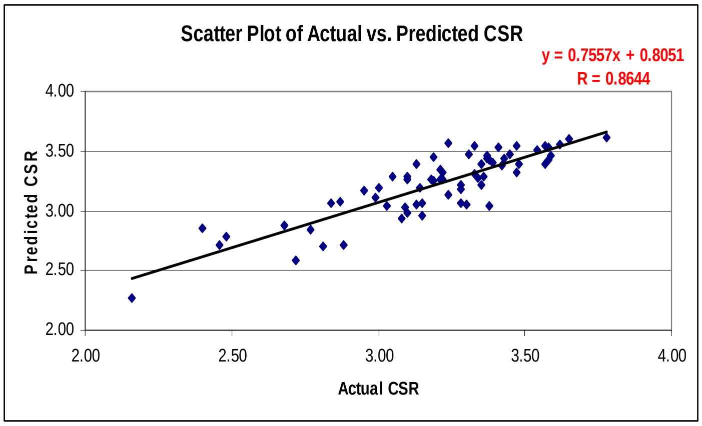

Figure 5. Scatter Plot for Top Two Useful Variables

6. CONCLUSIONS AND FUTURE RESEARCH

MTS was applied to the six parameters employed for the ride assessment to determine the useful variables. The discriminating ability of MTS was clearly shown through the proper selection of two groups, normal and abnormal. Then using an orthogonal array and signal-to-noise ratio, the useful variables were determined. The number of useful vehicle attributes was reduced from six factors to five factors.

The correlation of the two useful variables to the measured consumer satisfaction rating was 0.864 which equaled the correlation using a neural network. The neural network technique used a back propagation algorithm with an input layer, one hidden layer and an output layer. However, neural networks required considerably more time and data to achieve this level of correlation. Another key benefit was the required data for constructing the normal set and test group was considerably less than required for the neural network analysis. When compared to neural networks, MTS produces better accuracy using less data with equivalent correlation.

The methodology outlined in this research will enable producers to evaluate consumer satisfaction and accordingly apply that knowledge to the designing of the attributes delivered in a product. This research will provide a method for understanding and, therefore, meeting consumer requirements. Consequently, this will result in higher consumer satisfaction and lower costs by employing only those variables that are key to the achievement of consumer satisfaction ratings. Furthermore, the relationship developed will enable producers to quantify the impact of changes in part level components on consumer satisfaction.

Further research should be conducted to utilize the methodology developed in other contexts and industries that drive the requirements down throughout all levels of a system. The analysis performed relates subsystem levels to consumer satisfaction. Future research should incorporate subsystem, sub-subsystem and piece part levels to fully gain the relationship between these lower level components and overall consumer satisfaction. This will enable the producer quantify the impact on consumer satisfaction for actions such as changing suppliers for a particular component or modifying design specification. The methodology proposed will provide information for making business decisions.

A key aspect that should be incorporated into this analysis is cost data. Important business decisions are made on the basis of cost and consumer impact. System optimization would be based on equations using cost data and constraints in the feasible region for component specifications. The objective function would be to minimize the cost and maximize consumer satisfaction.

REFERENCES

[1] Asada M. (2001), Wafer yield prediction by the Mahalanobis-Taguchi system; *IEEE International Workshop on Statistical Methodology* 6; 25-28.

[2] Fowlkes W.Y., Creveling C.M. (1995), Engineering methods for robust product design; Addison-Wesley, Reading, MA.

[3] Hong J., Cudney E.A., Taguchi G., Jugulum R., Paryani K., Ragsdell K. (2005), A comparison study of Mahalanobis-Taguchi system and neural network for multivariate pattern recognition, *ASME IMECE*, Orlando, FL.

[4] Jugulum R., Monplaisir L. (2002), Comparison between Mahalanobis-Taguchi system and artificial neural networks; *Journal of Quality Engineering Society* 10(1); 60-73.

[5] Kiemele M.J., Schmidt S.R., Berdine R.J. (1999), Basic statistics: Tools for continuous improvement; Air Academy Press, Colorado Springs, CO.

[6] Lande U., Mahalanobis distance: A theoretical and practical approach, http://biologi.uio.no/fellesavdelinger/finse/spatialstats/Mahalanobis%20distance.ppt, July 11, 2004.

[7] Manly B.F.J. (1994), Multivariate statistical methods: A primer; Chapman & Hall, London.

[8] Taguchi G., Jugulum R. (2002), The Mahalanobis-Taguchi strategy; John Wiley & Sons, Inc., New York.

[9] Wu Y. (2004), Pattern recognition using Mahalanobis distance; *Journal of Quality Engineering Forum* 12 (5); 787-795.