Journal of Industrial and Systems Engineering
Vol. 1, No. 4, pp 281-292
Winter 2008

Identifying Useful Variables for Vehicle Braking Using the Adjoint Matrix Approach to the Mahalanobis-Taguchi System

Elizabeth A. Cudney[1], Kioumars Paryani[2*], 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 (MTS) is a diagnosis and forecasting method for multivariate data. Mahalanobis distance (MD) 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. MTS 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 Adjoint Matrix Approach to MTS for vehicle braking to identify a reduced set of useful variables in multidimensional systems.

Keywords: Mahalanobis-Taguchi system (MTS), Mahalanobis distance (MD), Adjoint matrix, Pattern recognition, Orthogonal array (OA), Signal-to-noise ratio (SN), Mahalanobis space (reference group).

1. INTRODUCTION

The purpose of this research is to develop a relationship between vehicle attributes and measured customer satisfaction ratings 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 Mahalanobis distance (MD) values (Taguchi and Jugulum, 2002).

Consumers assess and perceive 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 needs by providing affordable products. Consumers evaluate vehicle attributes such as ride, handling, roominess, braking and acceleration. These vehicle level

* Corresponding Author

attributes are influenced by factors at all levels of the vehicle architecture, and these factors are often correlated. The goal of this research is to efficiently forecast consumer satisfaction measured as a function of available vehicle level performance data. This paper presents the application of the Adjoint Matrix Approach to MTS (MDA) technique to identify a reduced set of useful variables in multidimensional systems.

MTS can be used to minimize the number of variables required to forecast the performance of a system. The steps of the MTS analysis for vehicle braking are outlined in this paper. The objective of the MTS analysis for vehicle brake pedal feel is to explain the relationship between the brake parameters and measured consumer satisfaction rating (CSR). The purpose is to provide an understanding of the relationship and provide an opportunity to efficiently improve consumer driven quality.

2. REVIEW OF RELEVANT LITERATURE

Considerable research is available utilizing Mahalanobis distance for determining the similarities of values from known and unknown samples. However, existing research uses Mahalanobis distance to consider measurable and distinct variables. Consumer requirements or inputs to the research were concrete and clearly defined. However, this research will create a methodology to translate lower level characteristics using Mahalanobis distance into performance specifications and capabilities.

Mahalanobis distance is utilized to determine the misclassification of samples in research by Shen et al (2003). Two processes for table production are evaluated in this research including wet granulation and direct compression which are the two main methods used in tablet preparation. Pyrolysis-gas chromatography-mass spectrometry was used to discriminate the two processes. First, samples were removed that did not contain at least half the number of samples for one of the two classes. Principal components analysis (PCA) was then employed to determine the main principal components. Based on the PCA analysis, three factors and one sample were excluded. The data was further processed using Fisher discriminant analysis to classify the sample using unsupervised and supervised classification methods (Shen et al, 2003). Fisher discriminant analysis is an approach that is commonly used for feature extraction between groups. The results were evaluated using Mahalanobis distance to determine the misclassification rate.

Taguchi utilized the Mahalanobis-Taguchi System for diagnosis and pattern recognition. His research discussed a case study involving liver disease diagnosis in Tokyo, Japan using fifteen variables. The normal group consisted of 200 people that were determined to not have liver disease. For each abnormal group, he determined which variables were significant (Taguchi, 2000). In his research, Taguchi developed an eight-step procedure titled "Mahalanobis Distance for Diagnosis and Pattern Recognition System Optimization Procedure".

Garcia-Lagos et al. utilized Mahalanobis distance in a topology assessment for power systems. The research developed system architecture for use in a state estimator which worked with a bus-branch oriented network model. The architecture was developed in two stages including a preprocessing stage and a classification stage. The preprocessing stage transformed the measurements into output vectors for grid topologies. In the second stage, classification is determined using a layer of Gaussian potential function units based on the Mahalanobis distance (Garcia-Lagos et al, 2003). The units are established through input from the preprocessed vector. This indicates the degree to which a vector belongs to a specific topology. This assessment enabled the researchers to identify the actual topology.

Al-Otum proposes two algorithms for color morphology. The first is a Mahalanobis-color-distance-based morphological ordering algorithm. The second is a corrected component wise morphological ordering algorithm. Both algorithms employ a Mahalanobis color measure based on reduced and conditional ordering of the data to perform four basic morphological operators: dilation, erosion, opening, and closing. The data ordering is performed using the Mahalanobis color distance (MCD) in which the angle-valued pixels are replaced by a scalar (Al-Otuml, 2003). The two algorithms were evaluated using a perceptual image quality assessment.

Mahalanobis distance was used to maximize productivity in a new manufacturing control system by Hayashi et al. The research used Mahalanobis distance as a core to their manufacturing control system because of the method's ability to recognize patterns. Due to a very high utilization of certain processes, one of the manufacturing areas for wafers was experiencing a bottleneck. Approximately 50% of the primary tools in this manufacturing area had over 90% utilization. Mahalanobis distance was used in conjunction with manufacturing administrators to focus on routes that involve the shortest cycle time and reduce labor costs (Hayashi et all, 2001). The new system detected deviations from normal productivity much earlier, enabled root cause identification and prioritized resolution.

Asada used the Mahalanobis-Taguchi System to forecast the yield of wafers. Yield of wafers is determined by the variability of electrical characteristics and dust. The research focused on one wafer product with a high yield. Mahalanobis distances were calculated on various wafers to compare the relationship between yield and distance (Asada, 2001). The signal-to-noise ratios were used to indicate the capability of forecasting and the effect of the parameters.

A method for handling uncertainty in feature position is proposed in the work of Anandan and Irani. The uncertainty in feature position for 3D shape and motion in multiple views depends on the underlying spatial intensity structure. The purpose of this research is to minimize the covariance-weighted squared error which is the Mahalanobis distance. The raw data is transformed into a covariance-weighted space in which the noise is uncorrelated and equally distributed. Singular Value Decomposition (SVD) is then employed to minimize the objective function in the transformed data space. A second step is then performed to recover the shape and motion of the original data space (Anandan and Irani, 2002). Results are provided for varying degrees of uncertainty.

Pattern recognition using Mahalanobis distance was demonstrated in the work of Wu. In this research, pattern recognition was used to diagnose human health. The results of tests from a regular physical check-up were used as the characteristics. The correlation between different tests was shown. Mahalanobis distance was used to summarize the multi-dimensional characteristics into one scale. In this research the base point was difficult to define because it was a healthy person. People who were judged to be healthy for the past two years were considered to be healthy (Wu, 2004). The research considered diagnosis of liver function with the objective of forecasting serious disease until the next check-up. The approach provided a more efficient method and also avoided inhuman treatment which was previously used in double blind tests.

The research presented here is distinct from existing literature due to the nature of value perception and the direction of the relationship. The focus is on consumer satisfaction which is consumer specific and not definitive. Using consumer satisfaction ratings and vehicle performance data, the useful variables are determined. The research takes MTS one step further to forecast consumer satisfaction ratings using regression analysis. In previous literature, the flow of the relationship is distinct. Multicollinearity between factors is also not addressed in the literature. Based on the regression equations, the impact on consumer satisfaction from changes in lower level components can be determined.

3. MULTIDIMENSIONAL SYSTEMS

A system, by its nature, is multidimensional. In order to improve or optimize a system, it is necessary to understand the variables and noise conditions that affect the system performance. All variables do not affect system performance to an equal degree. Therefore, it is paramount to identify the critical sets of variables through discriminant analysis of the system. The critical sets of variables can then be used for diagnosis and prediction.

Generally speaking, a multidimensional system consist of an input signal, M, that is typically given by the consumer. Noise conditions arise from changes in consumer use of the product, wear and deterioration, and manufacturing variation. The control factors are used to provide information to make a decision about the system. The output, y, should ideally closely match the input signal and be a measurable characteristic. A graphical depiction of a generalized multidimensional system is given in Figure 1.

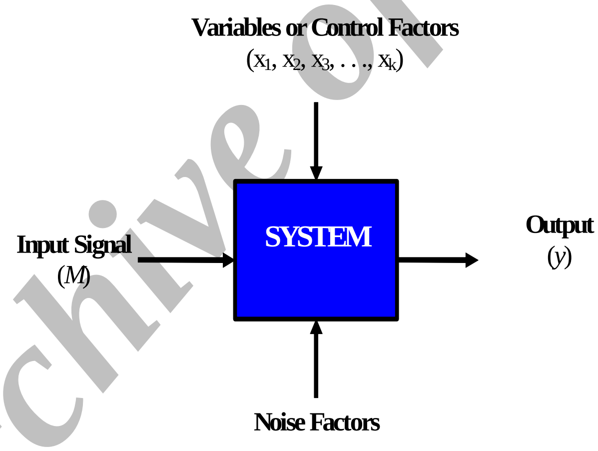

Figure 1. Multidimensional System

4. 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 approach (Taguchi 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 different from Euclidean distance because it addresses the correlations or distribution of the data points. 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}
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 (normal group) to that of an unknown set of values (abnormal group). MD has successfully been applied to a broad range of cases mainly because it is very sensitive to inter-variable changes in data. Also, because 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.

5. 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 (Taguchi 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 the wrong decision about the patterns if each variable is looked at 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 the 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 the MS is established, the number of attributes is reduced using 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 including and excluding that attribute to and from the system.

6. ADJOINT MATRIX APPROACH

An issue that must be addressed during the course of MTS analysis is multicollinearity. Multicollinearity occurs when there are strong correlations among three or more factors. In the presence of multicollinearity, the determinant of the correlation matrix will approach zero (Manly, 1994). Consequently, the resulting inverse matrix and scaled Mahalanobis distances (MD) will be difficult to compute; and they will be inaccurate if standard approaches are employed. The adjoint matrix approach (MDA) can be employed to handle multicollinearity.

In MDA the adjoint of the correlation matrix is used, instead of the inverse matrix, to calculate the distances. This approach can be used even when collinearity is not present. The adjoint of a square matrix *A* is obtained by replacing each element of *A* with its own cofactor and transposing the result. To distinguish from MD, the symbol MD_A is used to denote the distances obtained from an adjoint matrix. The relationship between the adjoint, determinant, and inverse is shown in the following equations.

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

$$MD_{Aj} = \frac{1}{k} Z_{ij}^T C_{adj} Z_{ij} \tag{3}$$

$$MD_j = \frac{1}{\det . C} MD_{Aj} \tag{4}$$

MD_A is similar to Mahalanobis distance (MD); however, there are different properties. MD and MD_A both represent the distances from the healthy or normal group and can be used to measure the degree of abnormality. In the adjoint matrix approach, the Mahalanobis space (MS) also contains means, standard deviations, and correlation structure of the healthy or normal group. However, here the MS is not called the unit space since the average of the MD_A 's is not unity.

7. MDA ANALYSIS OF BRAKE PEDAL FEEL

A total of 76 data sets were provided for analysis. However, 51 of these data sets had missing data for at least one of the brake pedal feel parameters. The remaining 25 good data sets were utilized for the MTS analysis.

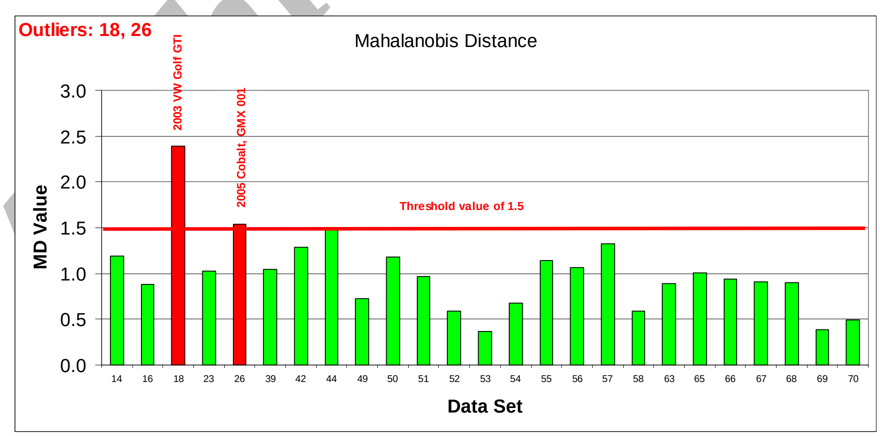

Figure 2. MD Values and Corresponding Outliers for Brake Pedal Feel Indices

The MD was calculated for the 25 data points using the 9 brake pedal feel index parameters. Using Mahalanobis distance, two outliers were identified and removed from the brake pedal feel indices as the first step in the MTS analysis. .Figure 2 illustrates the MD values for the 25 data points and the outliers. Using an arbitrary distance threshold value for MD of 1.5, two data sets are determined as outliers, i.e., 18, and 26.Minitab's matrix plot was utilized as an alternative and supplementary approach for identifying the outliers. The matrix plot was constructed for the 25 data points using the nine brake pedal feel index parameters as shown in Figure 3.

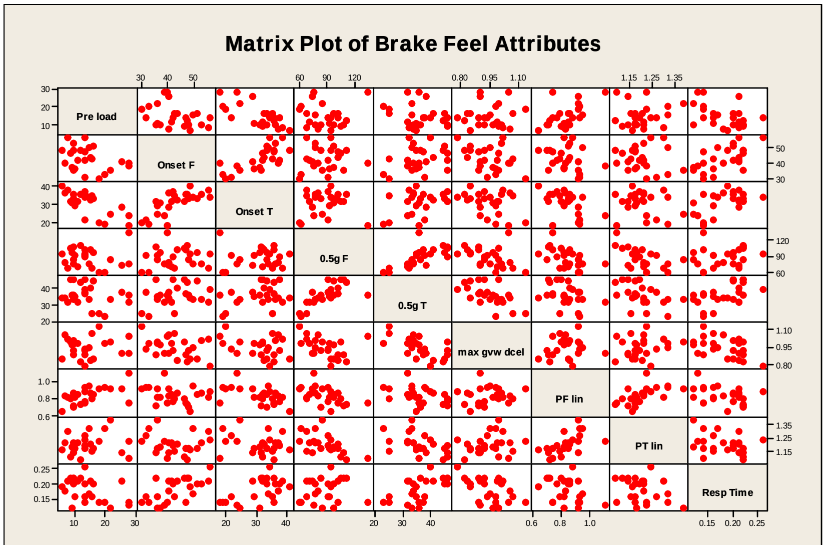

Figure 3. Matrix Plot of Brake Pedal Feel Indices for Identifying Outliers

Outliers can be selected using the matrix plot by visual inspection of these plots. No outliers were recognized for the brake pedal feel indices using the matrix plot. Using the MD threshold method two outliers were selected; whereas the matrix plot approach identified no outliers. Consequently, as mentioned earlier, two outliers were removed utilizing both the MD and the matrix plot.
The remaining 23 data sets were then used to perform the brake pedal feel analysis using the adjoint matrix approach (MDA). Let us emphasis again that the MDA approach is utilized in this analysis due to its ability to effectively handle multicollinearity. Nine brake pedal feel indices are used in the analysis.

Table 1. MD Values for the Normal Group

Data Set ID	CSR	MD_A
67	3.52	0.678
63	3.56	0.956
42	3.60	0.938
44	3.61	0.872
51	3.62	0.984
52	3.62	0.980
53	3.62	0.752
58	3.62	0.990
65	3.62	0.725
66	3.62	0.977
69	3.62	0.975

In implementing MDA, the first step is to define a “normal” or “healthy” group as the Mahalanobis space (MS). The Mahalanobis distance was then calculated using the data on the nine indices 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 (CSR) rating.

The MDs were then calculated for the test group. Table 2 summarizes these MDs. The data sets are again listed in order of ascending CSR.

Table 2. MD Values for the Test Group

Data Set ID	CSR	MD_A
16	3.45	9.665
49	3.46	15.952
54	3.46	23.692
57	3.46	48.764
56	3.70	18.884
23	4.00	77.213

A summary of the MD values for both the normal group and test group is provided in Figure 4.

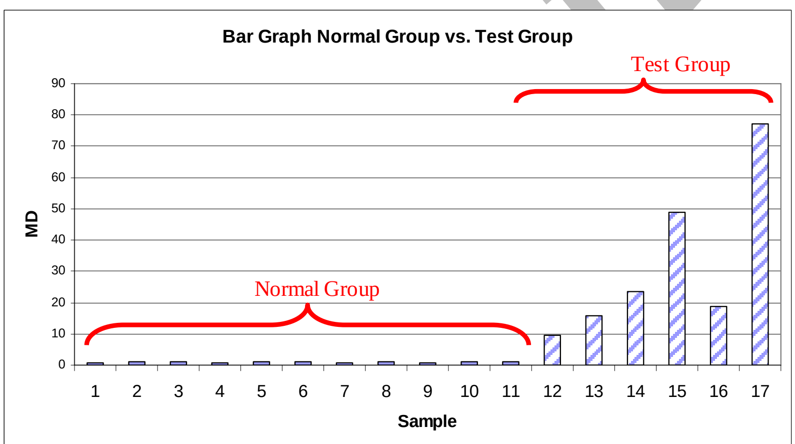

Figure 4. MD Values for the Normal Group and Test Group

The next step of the analysis is to optimize the discriminating ability of the MTS system using orthogonal arrays and signal-to-noise ratio. Table 3 shows the L_{16} orthogonal array used for the analysis of the nine brake pedal feel indices.

For each experiment in the OA a signal-to-noise (S/N) ratio is calculated. The S/N ratios, obtained from the test MDs, are used as the response for each combination of OA. The useful set of variables is determined by evaluating the "gain" in S/N ratio. The gains for the nine brake pedal feel indices are shown in Figure 5.

The next step in the analysis involves performing step-wise regression analysis. Variables are added to the regression analysis in descending order of the gains magnitude starting with the highest gain in S/N ratio. The regression equations shown in Table 4 were generated as each of the nine factors are systematically added to the regression analysis.

A summary of the correlation values as variables are added is shown in Table 5. Using the regression equations, the predicted consumer satisfaction values are calculated. As graphically shown in Figure 6, these predicted values are compared to the actual consumer satisfaction values for the five useful variables.

Table 3. L_{16} Orthogonal Array for the Nine Brake Pedal Feel Indices

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	2	2	1	2	1	1	2	2	*1*
2	2	2	1	2	1	1	2	1	*2*
3	2	2	1	1	2	2	1	2	*1*
4	2	2	1	1	2	2	1	1	*2*
5	2	1	2	2	1	2	1	2	*1*
6	2	1	2	2	1	2	1	1	2
7	2	1	2	1	2	1	2	2	1
8	2	1	2	1	2	1	2	1	2
9	1	2	2	2	2	1	1	2	2
10	1	2	2	2	2	1	1	1	1
11	1	2	2	1	1	2	2	2	2
12	1	2	2	1	1	2	2	1	1
13	1	1	1	2	2	2	2	2	2
14	1	1	1	2	2	2	2	1	1
15	1	1	1	1	1	1	1	2	2
16	1	1	1	1	1	1	1	1	1

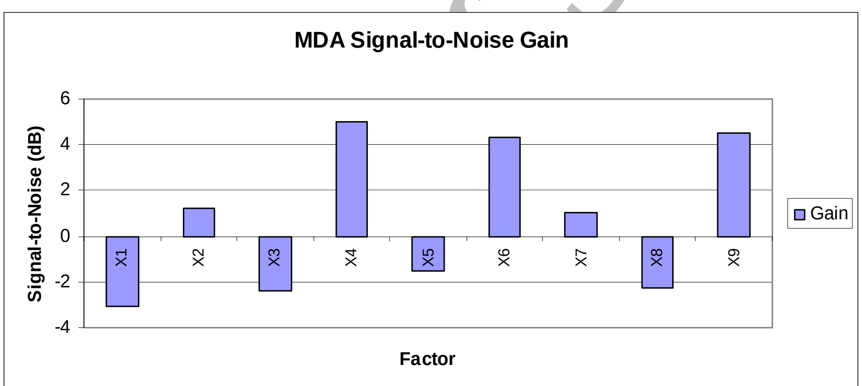

Figure 5. Calculated Gains for the Nine Brake Pedal Feel Indices

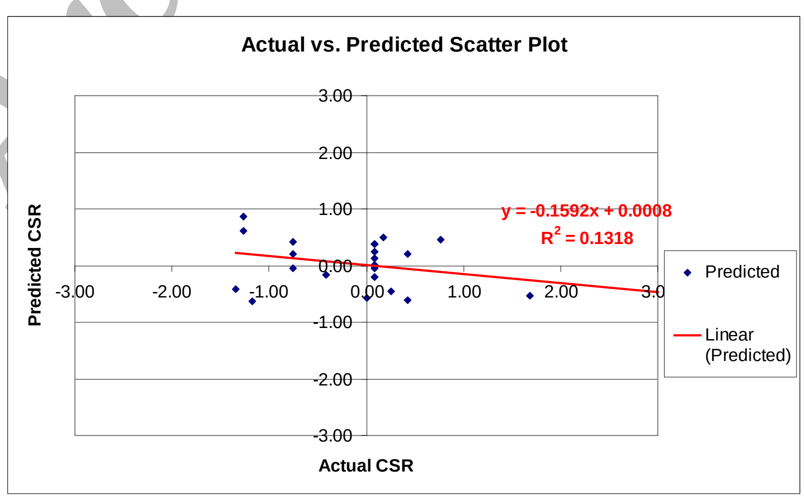

Figure 6. Measured Versus Predicted CS for the Useful Variables

Table 4. Step-wise Regression Analysis for the Useful Variables

Regression Equation	R-squared Value	R-Value
$CSR = -0.045 - 0.600[X_4]$	0.266	0.516
$CSR = -0.002 - 0.508[X_4] - 0.271[X_9]$	0.324	0.569
$CSR = 0.060 - 0.145[X_4] - 0.115[X_9] + 0.552[X_6]$	0.478	0.691
$CSR = 0.057 - 0.154[X_4] - 0.079[X_9] + 0.539[X_6] - 0.059[X_2]$	0.480	0.693
$CSR = 0.054 - 0.181[X_4] - 0.088[X_9] + 0.524[X_6] - 0.065[X_2] - 0.033[X_7]$	0.481	0.694
$CSR = 0.038 - 0.394[X_4] - 0.189[X_9] + 0.506[X_6] - 0.053[X_2] - 0.076[X_7] - 0.041[X_5]$	0.487	0.698
$CSR = 0.026 - 0.444[X_4] - 0.205[X_9] + 0.490[X_6] - 0.033[X_2] - 0.041[X_7] + 0.192[X_5] - 0.127[X_8]$	0.493	0.702
$CSR = 0.074 - 0.515[X_4] - 0.286[X_9] + 0.545[X_6] + 0.271[X_2] - 0.247[X_7] + 0.303[X_5] - 0.147[X_8] - 0.465[X_3]$	0.526	0.725
$CSR = 0.097 - 0.425[X_4] - 0.255[X_9] + 0.596[X_6] + 0.268[X_2] - 0.281[X_7] + 0.274[X_5] - 0.099[X_8] - 0.444[X_3] + 0.082[X_1]$	0.527	0.726

Scatter plots were then created to show the correlation between the measured versus predicted consumer satisfaction values. As shown in Figure 7, a scatter plot was created for the five useful variables. The scatter plot is useful to show the linearity of the prediction model.

Table 5. Regression Summary

Factor Added	R	Δ
X_4	0.516	
X_9	0.569	0.053
X_6	0.691	0.122
X_2	0.693	0.002
X_7	0.694	0.001
X_5	0.698	0.004
X_8	0.702	0.004
X_3	0.725	0.023
X_1	0.726	0.001

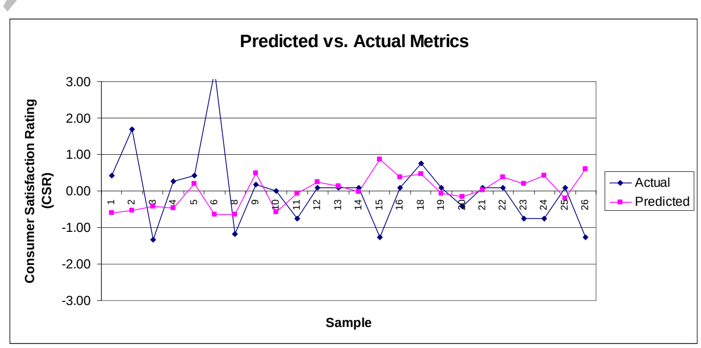

Figure 7. Scatter Plot for the Five Useful Variables

8. CONCLUSION

The brake pedal feel vehicle attribute data was evaluated using the Mahalanobis-Taguchi System. Multicollinearity is addressed in MTS by using the adjoint matrix approach (MDA). There are nine brake pedal feel factors. Of the original nine brake pedal feel factors, five factors were determined to be useful variables using the adjoint matrix method. The correlation coefficient between the actual and predicted values is 0.69.

This application of MDA 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 providing only characteristics that are key to the consumer. The relationship developed will enable producers to quantify the impact of changes in components level on consumer satisfaction.

9. FUTURE RESEARCH

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 subsystems, sub-subsystems and components levels to fully gain the relationship between design changes made at 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 components specifications. The objective function would be to minimize the cost and maximize consumer satisfaction.

REFERENCES

[1] Al-Otum, H.M.(2003), Morphological operators for color image processing based on Mahalanobis distance measure, *Optical Engineering*, 42 (9); 2595-2606.

[2] Anandan, P., Irani M.(2002), Factorization with uncertainty, *International Journal of Computer Vision*, 49 (2-3); 101-116.

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

[4] Garcia-Lagos, F., Joya, G., Marin, F.J., Sandoval F. (2003), Modular power system topology assessment using Gaussian potential functions, IEEE Proceedings-Generation Transmission and Distribution, 150 (5); 635-640.

[5] Hayashi, S., Y. Tanaka, Kodama E. (2001), A new manufacturing control system using Mahalanobis distance for maximizing productivity, *IEEE Transactions*, 15 (4); 59-62.

[6] Manly, B.F.J. (1994), Multivariate Statistical Methods: A Primer; Chapman & Hall, London.

[7] Shen, H., Carter, J.F., Brereton, R.G., Eckers C. (2003), Discrimination between tablet production methods using pyrolysis-gas chromatography-mass spectrometry and pattern recognition, *Analyst,* 128(3); 287-292.

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

[9] Taguchi, S.(2000), Mahalanobis-Taguchi system, ASI Taguchi Symposium, Detroit, MI.

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