Development, Analysis and Applications of a Quantitative Methodology for Assessing Customer Satisfaction Using Evolutionary Optimization

Sunith Bandaru[a], Abhinav Gaur[a], Kalyanmoy Deb[a],
Vineet Khare[b], Rahul Chougule[c]

[a]*Kanpur Genetic Algorithms Laboratory, Indian Institute of Technology Kanpur, Uttar Pradesh, India*
{sunithb,abnvgaur,deb}@iitk.ac.in
[b]*Amazon Development Centre (India) Pvt. Ltd., Bangalore, India*
vkhare@amazon.com
[c]*India Science Lab, General Motors Global R&D, Bangalore, India*
rahul.chougule@gm.com

KanGAL Report Number 2012012

Abstract

In this paper, we develop a method for the quantitative modeling of a Customer Satisfaction Index (CSI) function for consumer vehicles. The mathematical model is evolved using an evolutionary computation technique such that the satisfied and dissatisfied customers are equally distributed on either side of the mean satisfaction level. Instead of relying on a conventional survey based assessment, we extract various important features from the service (field failure) data of five different vehicle models and build optimized CSI functions for each. Next, the approach is extended so that a single CSI function can predict the satisfaction for all customers of all five models, thus providing a measure of the market's perceived quality of one vehicle model relative to another. Different combinations of vehicle models are used and the corresponding CSI functions are validated against the ratings published by Consumer Reports. A sensitivity analysis reveals interesting information about the features extracted from the service data. Thereafter, the CSI function which best differentiates between all five vehicle models is chosen for further use. Finally, we present two applications which use the chosen

CSI function to (i) identify high-priority problems for different vehicle models and (ii) data-mine the service data, both of which are very important from Customer Relationship Management (CRM) point of view.

Keywords: Customer Satisfaction Index (CSI), quantitative modeling, evolutionary optimization, Customer Relationship Management (CRM)

1. Introduction

Customer satisfaction has been defined in literature as "the state of mind that customers have about a company when their expectations have been met or exceeded over the lifetime of the product or service" [1, 2]. Customer satisfaction leads to customer retention, customer loyalty and product repurchase. Thus its measurement is an important and integral part of an effective Customer Relationship Management (CRM). Broadly speaking, satisfaction measures involve three psychological elements for evaluation of the product or service experience: (i) cognitive, which depends on the actual use of the product or service by the customer, (ii) affective, which depends on the customer's attitude towards the product or service or the company and (iii) behavioral, which depends on the customer's view regarding another product or service from the same company [3].

An important implication of the above definition for customer satisfaction is that it is subjective. Due to its non-quantifiability, most companies resort to a survey/questionnaire based assessment for the measurement of their products' perceived satisfaction. In this regard, years of research on customer behaviour has led to specification of ten domains of satisfaction, ranging from quality, efficiency, etc. to commitment to the customer and product innovation [4]. Surveys are designed to touch these domains. What actually to ask in the survey depends on the kind of product or service provided, the kinds of customers served, number of customers served, the longevity and frequency of customer/supplier interactions, and what is to be done with the results. The very nature of such surveys requires the customer to evaluate each statement on a psychometric scale (or a rating scale) [5]. Likert scales [6], semantic differential scales [7], smiling faces scales and percentage measures [8] are popular in that order. A typical five-level Likert item consists of a statement and asks the surveyee to choose among (a) Strongly disagree, (b) Disagree, (c) Neutral, (d) Agree, and (e) Strongly agree. Several studies exist which show the merit and demerits of different rating scales and how they should be interpreted.

In the automotive sector, OEMs depend on reports published by various marketing information firms like the American Customer Satisfaction Index [9], J.D.Power and Associates [10] and Consumer Reports [11] for gaining insights into vehicle quality. Quality analysis data provided are often focussed on questions related to number of failures in the field (e.g. Incidents Per Thousand Vehicles, IPTV [12] and Problems Per Hundred vehicles, PPH) for individual components like engine, transmission, etc. Limited emphasis is placed on the assessment of individual users' perception and satisfaction resulting from day to day use of the product. Moreover, survey based estimates rely on a small sample of the customers (around 200 to 400 per vehicle model [10]). Despite this, the surveys themselves are highly regarded and play a significant role in moulding the customer's attitude towards a particular vehicle model. To some extent, the surveys also help the OEMs in identifying major problem areas. The CRM policy of OEMs should therefore be flexible enough to take into account the information contained in these survey reports published annually.

For service based companies, Parasuraman et. al. [13, 14] proposed the 'gaps model' for estimating satisfaction objectively by using the gap between the customer's expectation and perceived experience of performance. Apart from these and a few other related studies, quantitative measurement of customer satisfaction has not received much attention in literature. The main reason is as follows: There are three practical approaches to measuring satisfaction, namely, post-purchase evaluation, periodic satisfaction surveys and continuous satisfaction tracking. Post-purchase evaluation (known as Initial Quality Study in the automotive sector) deals with satisfaction assessment *shortly after* the delivery of product or service. Periodic satisfaction surveys provide occasional snapshots of customer perceptions. Continuous satisfaction tracking is much like post-purchase evaluation but carried out over time. Post-purchase evaluations are very common and seems to be used across all sectors. Most products and services are however, not amenable to periodic assessment and therefore not enough data is obtained for a single customer to warrant a quantitative study.

Automotive OEMs, on the other hand, provide customers a warranty period which covers repairs and mechanical faults as part of the sale. Claims can be made by the customers at authorized dealerships and service stations which keep customer-specific records of these claims. Warranty data consists of claims data and supplementary data. A review of warranty data analysis methods for identifying early warnings of abnormalities in products, providing useful information about failure modes to aid design modification, estimating product reliability for deciding on warranty policy, and forecasting future warranty claims needed for preparing fiscal plans can be found here [15]. In this paper, we go a step further and use the same warranty data for obtaining a mathematical model for predicting customer satisfaction. Typically, customer satisfaction is measured at the individual level, but it is almost always reported at an aggregate level. We ensure that this is true for our model by employing a bottom-up approach to modeling.

Table 1 shows some basic statistics of the five vehicle models considered in this study. The numbers correspond to the warranty data of all vehicles serviced between January '08 and August '09. For anonymity, the total number of customers and claims for each model are shown relative to those of Model 5. The last row of the table shows the number of unique field failures, a common constituent of IPTV or PPH figures, that occurred in a vehicle model during the said period. A field failure refers to any vehicle related problem faced by the customer for which he/she had to visit a dealer or service station. Each unique field failure is associated with a corresponding repair code for classification purposes. Given the limited resources available with customer relation managers, it is only prudent to prioritize these field failures for root cause analysis and possible reduction. The methodology presented in this paper allows one to prioritize these unique field failures based on a quantitative measure of their potential for improvement in the customer's perception and hence the CSI. The method begins with the building of a quantitative model of the customer satisfaction index using an evolutionary optimization technique. The present work suggests an improvement to the method in [16] and validates it against Consumer Reports ratings of the vehicle models. The resulting CSI modeling function is then used to obtain the CSI Improvement Potential (CIP) for different types of field failures.

The rest of the paper is organized as follows. In Section 2 we describe the components of the dataset being used. Processing of this dataset and extraction of relevant features is described in Section 3. Section 4 describes the framework for the proposed satisfaction model built on these extracted features. Section 5 presents the methodology for obtaining the satisfaction model for a given vehicle model and Section 6 presents its extension to obtain a generalized satisfaction model when multiple vehicle models are involved. The results and their validation are presented in Section 7. A sensitivity anal-

Table 1: Some basic statistics of the vehicle models relative to Model 5

	Vehicle Model Number				
	1	2	3	4	5
Segment	Compact	Midsize	Luxury	Midsize	Luxury
Total Customers	19.61 × C	13.35 × C	1.93 × C	30.19 × C	C
Total Claims	11.10 × K	6.99 × K	1.57 × K	17.67 × K	K
Total Field Failures	1026	1084	776	1228	606

ysis is also performed on the obtained CSI function. Next, Section 8 presents two applications of the quantitative CSI model, which are very relevant from the CRM point of view. Section 9 concludes the paper.

2. Vehicle Sales and Service Data

As discussed above, warranty data consists of claims data and supplementary data. The datasets used in this study also contain service repairs made beyond the warranty period and hence we are referring to claims data as service data. Supplementary data does not change with time. In our case, the sales data serves as this component.

Before going any further, it would be beneficial to know what attributes of vehicles contribute to satisfaction. A recent study by J.D. Power and Associates lists the following as critical elements of customer satisfaction [17]:

1. Quality and reliability (24%) with respect to problems experienced with the vehicle across different vehicle sub-systems.
2. Vehicle appeal (37%) which includes design, comfort, features, etc.
3. Ownership costs (22%) which includes fuel consumption, insurance and costs of post-warranty service or repair.
4. Dealer service satisfaction (17%) with respect to service quality, service time, etc.

It can be seen that, satisfaction related to quality, reliability and service contributes close to 50% towards the overall satisfaction (including some contribution from the ownership costs). In this paper, we are primarily focussed on assessing quality, reliability and service satisfaction as they can be quantified through the field failure information obtained from sales and

Table 2: Data fields in the sales and claims data of a vehicle and the notation used in this work.

Vehicle Sales Data (One-time entry)		Vehicle Service Data (For i-th claim of a single visit)	
Vehicle ID No.	VIN	Vehicle ID No.	VIN
Sale date	d_0	Repair Start date	d_i
Mileage at sale	m_0	Repair End date	e_i
		Mileage at repair	m_i
		Repair cost	c_i
		Repair code	r_i

service data. Table 2 shows typical fields found in the sales and service data of a vehicle that can be used to assess this satisfaction.

It is not directly clear how the data fields in Table 2 can be used. Therefore, we first extract certain characteristic features $\mathbf{x}$ from the combined sales and service data, so that the CSI for a particular customer/vehicle can be approximated by a mathematical function f,

$$CSI_{vehicle} = f(\mathbf{x}). \tag{1}$$

We call $CSI_{vehicle}$ the CSI model for predicting the customer satisfaction of a given vehicle model.

From the manufacturer's point of view, a CSI model (f) which can aggregate the views of all the customers in a deterministic way is beneficial for identifying vehicle models with high perceived quality and, more importantly, for highlighting the problem areas across different vehicle models. Taking cues from literature, we can impose certain conditions on the function f so that it achieves the above stated goal. It has been hypothesized that the frequency distribution of customer satisfaction is a theoretically continuous one that is typically skewed to the left [18]. Moreover, it is also widely believed that this distribution is often the convergence of (two or) three nearly normal distributions of (two or) three statistically differing populations - dissatisfied, satisfied and extremely satisfied customers [18]. The work presented in this paper is based on this hypothesis. The direct implication is that, given a vehicle model, most of the customers will have a similar overall perception (satisfaction) of the vehicle model. This results in a $CSI_{vehicle}$ distribution with low variance as shown in Figure 1. A function f which models the CSIs based on this assumption would also make more sense for averaging over all customers thus obtaining an overall satisfaction value for the vehicle model under consideration. This is what we mean by the bottom-up approach mentioned earlier. By aggregating the $CSI_{vehicle}$ values for all the vehicles (customers) corresponding to a model, a representative CSI value for that vehicle model can be obtained.

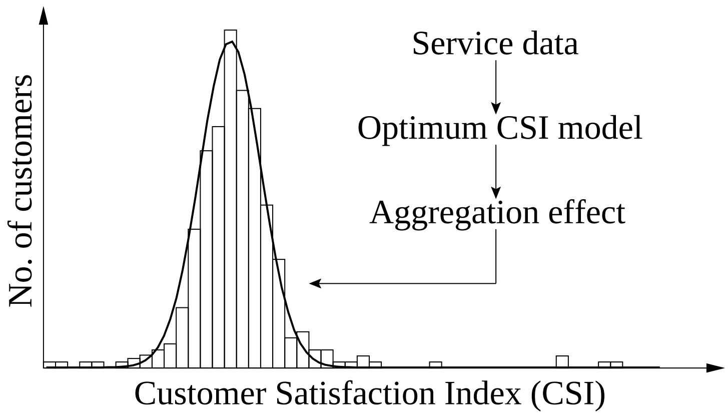


Figure 1: Expected frequency distribution of CSI from the proposed method which makes averaging possible.

3. Feature Extraction

The fields presented in Table 2, though representative of the customer satisfaction, cannot directly be used in a mathematical function. To have a computationally tractable method for modeling the CSI, we have identified six features that can be deterministically calculated for each vehicle using the combined sales and service data. The relation between these features and the satisfaction can also be logically established as shown below:

1. x_1: Number of visits made by a customer. The VIN or Vehicle Identification Number is unique for each vehicle. By counting the number of times a particular VIN occurs in the service dataset, the number of visits made by the customer owning that vehicle can be determined. More visits mean lower customer satisfaction and hence the dependency can be modeled as $CSI_{vehicle} \propto 1/x_1$.
2. x_2: Total number of days for which the vehicle was unavailable to the customer. The waiting time for a customer is the calender difference

between the repair start date and the delivery date. We sum these differences over the number of visits to get $x_2 = \sum_{i=1}^{x_1}(e_i - d_i)$. Longer waiting times have a negative impact on satisfaction which can be modeled as $CSI_{vehicle} \propto 1/x_2$.

3. x_3: Sum of all service/repair costs. These costs include the labor costs, part costs and miscellaneous costs, if any. The total expenditure on a vehicle for the given period can be obtained as $x_3 = \sum_{i=1}^{x_1} c_i$. And logically it follows that $CSI_{vehicle} \propto 1/x_3$.
4. x_4: Average time interval between visits. The time to first visit is the calendar difference between the earliest visit date in the service data and the vehicle sale date. Thereafter, the time interval between subsequent visits can be obtained from the service data alone. The cumulative time intervals are averaged over the number of visits. Mathematically,
$$x_4 = \frac{1}{x_1}\left(d_1 - d_0 + \sum_{i=2}^{x_1}(d_i - d_{i-1})\right). \tag{2}$$
A larger value of x_4 means longer problem-free vehicle use and hence higher customer satisfaction. Therefore, we have $CSI_{vehicle} \propto x_4$.
5. x_5: Average miles run between visits. Like the time intervals, the miles run by the vehicle without problems can be obtained from the odometer readings in the sales and service data as,
$$x_5 = \frac{1}{x_1}\left(m_1 - m_0 + \sum_{i=2}^{x_1}(m_i - m_{i-1})\right). \tag{3}$$
It is easy to conclude that $CSI_{vehicle} \propto x_5$.
6. x_6: Sum of problem severity ratings. Each vehicle visit is associated with a repair code r_i which defines the type of failure. All repair codes are assigned a severity rating between 1 (minor problem; e.g. oil change) and 5 (major problem; e.g. engine replacement) by domain experts. Since severity rating has a negative impact on the CSI, we have $CSI_{vehicle} \propto 1/x_6$.

4. CSI Model Framework

The characteristic features extracted above are still not directly usable since they vary in different ranges. Table 3 shows some statistics of the features extracted from the service data of the five vehicle models that we use in this work. It is apparent that the direct use of these features for constructing f will lead to a biased CSI model, since numerically they will have varying effect on the CSI. A normalization procedure is adopted to bring the features in the range $[0, 1]$ using,

$$x_i^{nr} = \frac{x_i - \min_j x_i^{(j)}}{\max_j x_i^{(j)} - \min_j x_i^{(j)}} \quad \forall\, i \in \{1, 2, \ldots, 6\},\ j \in \{1, 2, \ldots, C\}, \tag{4}$$

where C is the total number of customers of a vehicle model. It is easy

Table 3: Some statistics of extracted features

	Stat.	x_1	x_2	x_3	x_4	x_5	x_6
Model 1	Min.	1	1	0	1	1	1
	Max.	20	357	11253.76	429	799686	50
	Mean	1.39	1.46	155.93	110.63	5317.23	3.14
Model 2	Min.	1	1	0	1	1	1
	Max.	12	371	8195.94	395	115671	39
	Mean	1.32	1.37	199.07	137	7810.22	3.09
Model 3	Min.	1	1	1.91	1	1	1
	Max.	10	77	12188.07	408	999959	51
	Mean	1.67	1.74	349.01	114.57	5502.76	4.21
Model 4	Min.	1	1	0	1	1	1
	Max.	17	368	9985.60	433	333024	45
	Mean	1.45	1.59	148.10	129.33	6573.53	2.65
Model 5	Min.	1	1	5	1	1	1
	Max.	10	248	9134	459	78501	35
	Mean	2.01	2.35	458.71	118.70	5162.20	5.03

show that such a linear normalization does not effect the frequency distribution. Keeping in mind the logical dependencies established earlier, we now introduce the following feature transformation for simplicity of notation:

$$X_i = \frac{1}{(1 + x_i^{nr})} \quad \text{, for } i \in \{1, 2, 3, 6\}, \tag{5}$$

$$X_i = (1 + x_i^{nr}) \quad \text{, for } i \in \{4, 5\}. \tag{6}$$

In this work, we call X_i's as the *transformed features* in contrast to the x_i's introduced in the last section, which are called *extracted features*. The addition of the constant (one here) in Equation (5) and (6) ensures that the CSI does not approach infinity for vehicles with minimum corresponding x_i(s). Equation (1) can now be re-written as,

$$CSI_{vehicle} = f(X_1, X_2, X_3, X_4, X_5, X_6). \quad (7)$$

The functional form for f is not defined *a priori*. However, as stated earlier, our primary goal is to obtain a CSI model with low variance (as shown in Figure 1) for a given vehicle model. Secondly, the model should be *flexible* enough to differentiate between two or more vehicle models which have clearly different perceived quality in the market. This is our secondary goal. The modeling framework proposed here is an adaptive one, i.e. the mathematical form for f is not assumed but is adaptively built by an algorithm. The model is composed of six terms. Each term can be a product of different transformed features so that,

$$CSI_{vehicle} = \sum_{l=1}^{6} T_l, \text{ where } T_l = \prod_{i=1}^{6} X_i^{\alpha_{il}\beta_{il}} \quad \forall\, i, l \in \{1, \ldots, 6\}. \quad (8)$$

Here β_{il}'s are Boolean decision variables which decide the presence (when $\beta_{il} = 1$) or absence (when $\beta_{il} = 0$) of the i-th transformed feature in the l-th term. When $\beta_{il} = 1$, α_{il} denotes the corresponding exponent.

The six multiplicative terms can in turn be multiplied or added to each other. Again, a Boolean encoding decides which terms to multiply and which ones to add. We adopt 1 for '+' operator and 0 for '×' operator. These intermediate Boolean variables are represented by γ_l. The inverse operations (division and subtraction) are not considered since the dependency of $CSI_{vehicle}$ to each extracted feature x_i is already incorporated into the transformed features X_i's. For the same reason the powers α_{il}'s are considered to be non-negative real numbers lying between 0 and 1. With six terms, we have five additional Boolean decision variables. The 41 (= 36 β's+5 γ's) Boolean variables can be represented using a binary string.

Figure 2 shows an example of the representation scheme adopted in this paper. Every seventh bit is the variable γ_l which defines the arithmetic operation between the T_l-th and T_{l+1}-th terms. The usual order of precedence is followed for the operators (multiplication and then addition) to evaluate

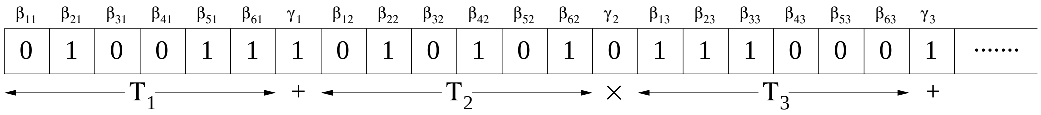


Figure 2: Binary representation for the proposed adaptive form of the CSI model.

the CSI. Whenever a term evaluates to unity (i.e. when $\beta_{il} = 0 \;\; \forall \; i$), it is simply ignored in further computations. The illustrated string evaluates to

$$\begin{aligned} CSI_{vehicle} &= T_1 + T_2 \times T_3 + \cdots \\ &= X_2^{\alpha_{21}} X_5^{\alpha_{51}} X_6^{\alpha_{61}} + X_2^{\alpha_{22}} X_4^{\alpha_{42}} X_6^{\alpha_{62}} \times X_1^{\alpha_{13}} X_2^{\alpha_{23}} X_3^{\alpha_{33}} + \cdots \\ &= X_2^{\alpha_{21}} X_5^{\alpha_{51}} X_6^{\alpha_{61}} + X_1^{\alpha_{13}} X_2^{\alpha_{22}+\alpha_{23}} X_3^{\alpha_{33}} X_4^{\alpha_{42}} X_6^{\alpha_{62}} + \cdots \end{aligned}$$

This modeling approach resembles a genetic programming module with $\mathcal{T} = \{X_i \;\forall\; i\}$ as the terminal set and $\mathcal{F} = \{\times, +\}$ as the functional set [19] without introducing the computational cost associated with such generic systems.

5. Customer Level CSI Model

For achieving a customer level CSI model we now focus on our primary goal of obtaining a CSI function which, when evaluated for different customers produces a distribution with low variance. A narrower CSI distribution means better agreement among the customers thus making averaging more sensible for obtaining the overall satisfaction. It is intuitive that this can be accomplished by minimizing the variance (or standard deviation) of the CSI values of all customers. However, there are two situations which can lead to trivial solutions.

Firstly, since the exponents α_{il}'s are non-negative, a simple minimization of the variance causes them to approach the lower bound of zero giving rise to severely right skewed CSI distributions after normalization. This is shown in Figure 4, described later. To avoid this, we incorporate a second 'helper objective' which minimizes the skewness of the distribution. Secondly, the presence of Boolean variables may cause $\beta_{il} = 0 \;\forall\; i, l$ giving a trivial distribution of zero variance. The following set of constraints are used to prevent this situation.

$$\sum_l \beta_{il} \geq 1 \quad \forall \; i, l$$

which basically impose that each extracted feature X_i is used in at least one of the terms. This also makes practical sense since the six extracted features are the only information that can be derived from the sales and service data and we would want the resulting CSI model to use all of them.

This optimization is carried out using a genetic algorithm (GA). The binary representation of the model discussed above was designed with this in mind. Equation (9) gives the bi-objective optimization problem formulation for finding the customer level CSI model from the sales and service data of C customers of a vehicle model.

$$\begin{array}{llcl}
\text{Minimize} & \sigma & & \\
\text{Minimize} & |g| & & \\
\text{Subject to} & \sum_l \beta_{il} \geq 1 & \forall & i \in \{1,\ldots,6\}, \\
\text{36 real variables:} & 0 \leq \alpha_{il} \leq 1 & \forall & i,l \in \{1,\ldots,6\}, \\
\text{36 Boolean variables:} & \beta_{il} \in \{0,1\} & \forall & i,l \in \{1,\ldots,6\}, \\
\text{5 Boolean variables:} & \gamma_l \in \{0,1\} & \forall & l \in \{1,\ldots,5\}, \text{where,}
\end{array} \tag{9}$$

$$\begin{aligned}
\sigma &= \sqrt{\frac{1}{C}\sum_{j=1}^{C}\left(CSI^{nr}_{vehicle,j} - \mu\right)^2}, \qquad \mu = \sum_{j=1}^{C} CSI^{nr}_{vehicle,j} \Big/ C, \\
g &= \frac{1}{C}\sum_{j=1}^{C}\left(CSI^{nr}_{vehicle,j} - \mu\right)^3 \Big/ \sigma^3.
\end{aligned}$$

where $CSI^{nr}_{vehicle,j}$ is the normalized value of $CSI_{vehicle}$ evaluated for the j-th customer. The normalization is linear between 0 and 1, performed using maximum and minimum values of $CSI_{vehicle,j}$. Note the arbitrary upper bound on α_{il}'s. Any other value may also be used. Figure 3 shows the non-dominated solutions obtained by solving (9) for each of the five vehicle models (refer Table 1) individually using NSGA-II [20]. All fronts are obtained using the following settings:

1. Population size: 2000,
2. Number of generations: 500,
3. Tournament selection with size 2,
4. Simulated binary crossover [21] with $p_c = 0.9$ and $\eta_c = 10$ for α's,
5. Polynomial mutation [22] $p_m = 0.05$ and $eta_m = 50$ for α's,
6. Single-point crossover with $p_c = 0.9$ for binary string (β's and γ's),

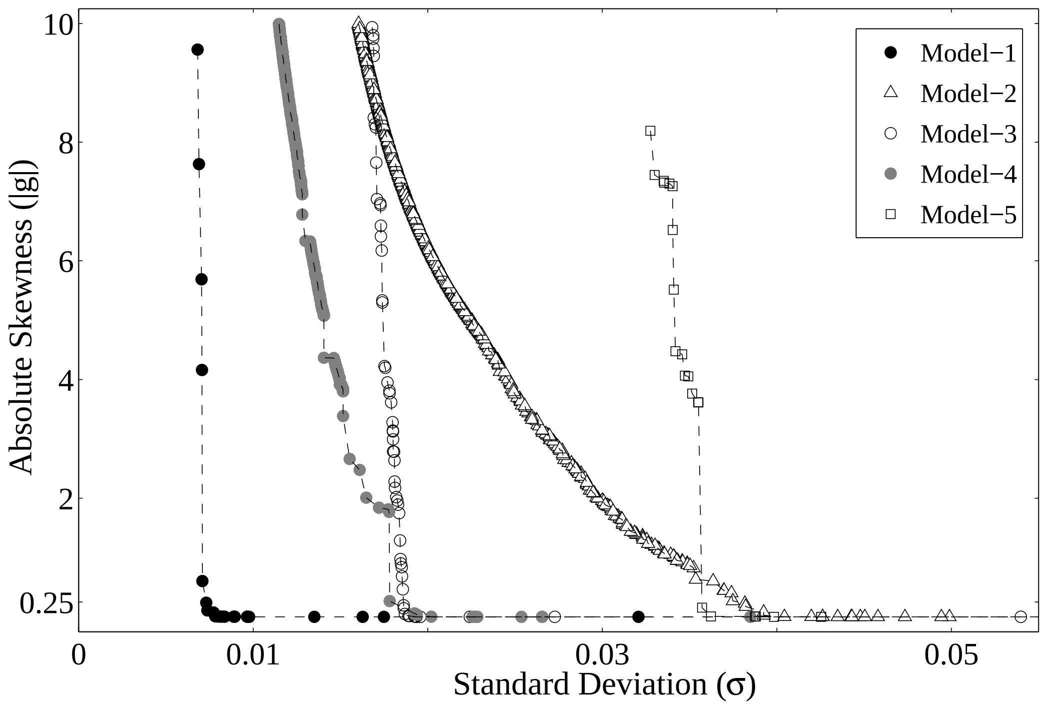


Figure 3: Non-dominated solution sets for the five vehicle models considered in this study, obtained by solving Equation (9). Note that each point shown on these trade-off fronts represents a different CSI model.

7. Bitwise mutation with $p_m = 0.15$ for binary string (β's and γ's).

Each trade-off solution in Figure 3 represents a different CSI model. The sharp kink in each of the trade-off fronts indicates the presence of a *knee*. The knee point of a two dimensional Pareto-optimal front is the solution which gives the best trade-off with respect to both objectives. Due to this characteristic it is also often called the *preferred solution*. Figure 4 shows the CSI distributions corresponding to three different points on the trade-off front of Model 4. Note that the leftmost solution is a trivial solution obtained when all α's approach their lower bounds as stated earlier. On the other hand, the rightmost solution yields a relatively wide CSI distribution which makes it difficult to estimate the overall CSI rating of the vehicle model from the CSI values of individual customers, because averaging in such cases is not meaningful. The second helper objective allowed many CSI models to coexist among which the knee solution is found to be most favourable.

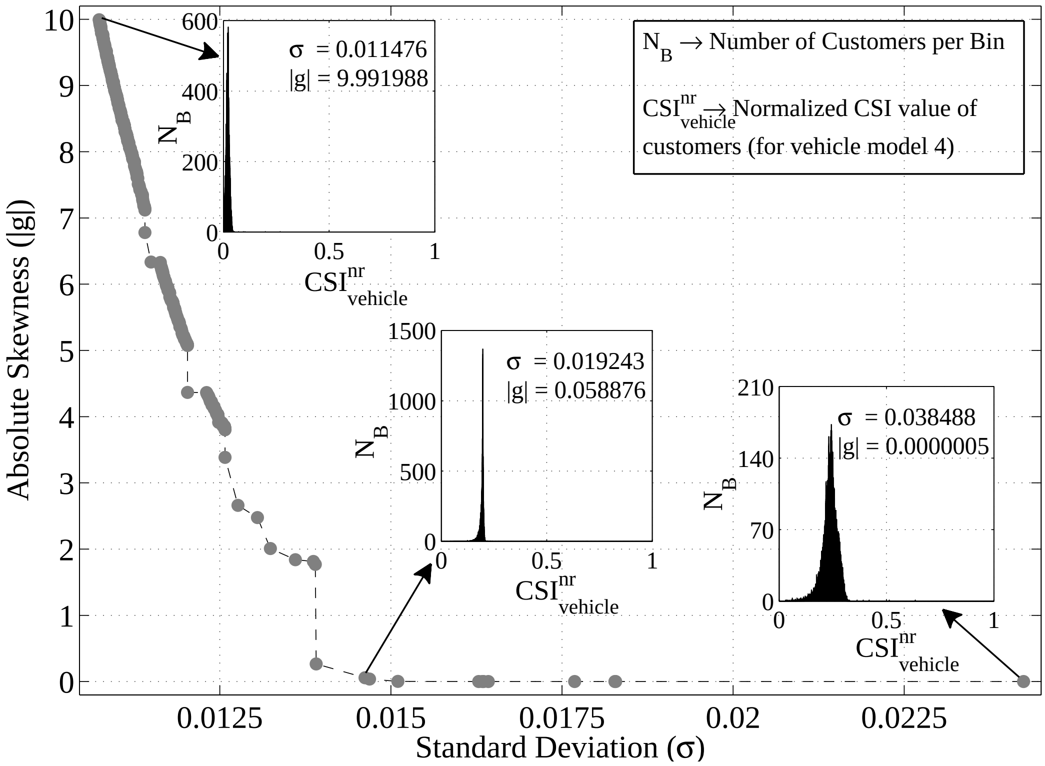


Figure 4: The figure shows the non-dominated solution set obtained from solving Equation (9) for Model 4. The three insets show the CSI distributions corresponding to, (a) the leftmost point of the trade-off front, (b) the *knee* point of the trade-off front, and, (c) the rightmost point of the trade-off front.

Other solutions in the knee region also yield similar distributions, making this region a potential search space for good CSI models. We use the bend-angle approach proposed in [23] to identify the knee solution. Table 4 shows the knee-point CSI models and their corresponding objective values for all five vehicle models obtained from the trade-off fronts shown in Figure 3.

6. Multiple Vehicle CSI Model

The customer level CSI modeling technique presented above fulfils our primary goal of obtaining a low variance and low skewness CSI distribution. However, as seen in Table 4 the obtained CSI models differ between vehicle models and hence do not provide a common basis for comparing or rank-

Table 4: Objective function values and the knee-point CSI models for all vehicle models.

		Knee-point solution details
Model 1	Objectives	$\sigma = 0.0071, \lvert g\rvert = 0.0021$
	$CSI_{vehicle}$	$T_1 + T_2 \times T_3 + T_4 + T_5 + T_6$ $= X_2^{0.0632} X_3^{0.0184} X_5^{0.0696} + X_2^{0.0632} X_5^{0.9175} + X_2^{0.5889} X_3^{0.0704}$ $+ X_1^{0.0153} X_2^{0.7322} X_6^{0.0375} + X_2^{0.9479} X_4^{0.0011}$
Model 2	Objectives	$\sigma = 0.0392, \lvert g\rvert = 0.0253$
	$CSI_{vehicle}$	$T_1 + T_2 \times T_3 \times T_4 \times T_5 + T_6$ $= X_3^{0.9238} X_4^{0.0092} X_5^{0.0048} + X_1^{0.3450} X_2^{0.8507} X_3^{0.9177} X_4^{0.0071}$ $+ X_2^{0.0292} X_3^{0.2830} X_5^{0.7601} + X_2^{0.1085} X_3^{0.8603}$
Model 3	Objectives	$\sigma = 0.0178, \lvert g\rvert = 0.2665$
	$CSI_{vehicle}$	$T_1 + T_2 \times T_3 \times T_4 \times T_5 + T_6$ $= X_1^{0.1089} X_2^{0.2060} X_3^{0.9139} + X_2^{0.4329} X_3^{0.9731}$
Model 4	Objectives	$\sigma = 0.0187, \lvert g\rvert = 0.0498$
	$CSI_{vehicle}$	$T_1 + T_2 \times T_3 + T_4 + T_5 + T_6$ $= X_2^{0.8242} X_6^{0.0259} + X_1^{0.0030} X_2^{0.9037} X_3^{0.0217} + X_5^{0.7632}$ $+ X_2^{0.2979} + X_3^{0.0283}$
Model 5	Objectives	$\sigma = 0.0362, \lvert g\rvert = 0.0072$
	$CSI_{vehicle}$	$T_1 + T_2 \times T_3 + T_4 + T_5 + T_6$ $= X_2^{0.7563} X_4^{0.0287} X_5^{0.0230} X_6^{0.0303} + X_2^{0.4977} X_5^{0.6763} + X_2^{0.8656}$ $+ X_2^{0.9701} X_6^{0.0142} + X_2^{0.8672} X_3^{0.0151}$

ing them. Our secondary goal, as stated earlier, is therefore to modify the problem formulation in (9) so that a single CSI model can be used to differentiate between two or more vehicle models as distinctly as possible. For this purpose, we propose to maximize the absolute difference between the average CSI values of the vehicle models under consideration. The results from customer level CSI models tell us that the best CSI models are obtained in the knee region. We can therefore constrain the search space to a region which contains the knee points of all five vehicle models by converting the variance and skewness objectives into additional constraints. From Figure 3, it is clear that the knee regions of all the five Pareto fronts corresponding to the customer level CSI models of the five vehicle models are bounded by

$\sigma \leq 0.05$ and $|g| \leq 0.25$. The single objective optimization problem proposed for multiple vehicle CSI modeling is,

$$
\begin{array}{lllll}
\text{Maximize} & \displaystyle\sum_{\{m,n|m\neq n\}} |\mu_m - \mu_n| & & & , \\
\text{Subject to} & \sigma_m \leq 0.05 & \forall & m \in \{1,2,\ldots,5\}, & \\
 & |g_m| \leq 0.25 & \forall & m \in \{1,2,\ldots,5\}, & \\
 & \sum_l \beta_{il} \geq 1 & \forall & i,l \in \{1,2,\ldots,6\}, & \\
\text{where,} & & & & \\
\text{36 real variables:} & 0 \leq \alpha_{il} \leq 1 & \forall & i,l \in \{1,2,\ldots,6\}, & \\
\text{36 Boolean variables:} & \beta_{il} \in \{0,1\} & \forall & i,l \in \{1,2,\ldots,6\}, & \\
\text{5 Boolean variables:} & \gamma_l \in \{0,1\} & \forall & l \in \{1,2,\ldots,5\}. &
\end{array}
\tag{10}
$$

Here m and n represent the indices of the vehicle models considered for obtaining the multiple vehicle CSI model. σ_m, μ_m and g_m are respectively the standard deviation, mean and skewness of the normalized CSI values of the m-th vehicle model (having C_m customers). For the five vehicle models, we can use any number and combination of them to obtain a corresponding multiple vehicle CSI model. Again, a GA-based optimization algorithm is used to solve the above problem. The sales and service data of all the vehicle models considered are given as input.

In the following discussion we refer to any combination of vehicle models as a Training Set (TS). For simplicity of notation, we represent a training set as TS{`List of vehicle models`} and the corresponding CSI model as, $CSI_{\{\texttt{List of vehicle models}\}}$. For example, the training set containing Models 2, 3 and 4 is TS{2,3,4} and its corresponding CSI model will be referred to as $CSI_{\{2,3,4\}}$.

7. Results and Discussion

The five vehicle models in this study have been chosen from different market segments as shown in Table 1. These models have been in the market for some time and survey based assessment is available for them. Unlike the American Customer Satisfaction Index and surveys by J.D. Power and Associates, Consumer Reports surveys assess customer satisfaction based on three aspects: performance, safety and reliability. In the context of the methods proposed in this paper, where we use service data for CSI modeling, Consumer Reports reliability ratings provide the most relevant assessment for comparisons. According to a recent Consumer Reports study [24], the order of reliability ratings of the five vehicle models is given by,

$$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_4 \approx CSI_2. \tag{11}$$

meaning that Model 5 has the worst overall customer perception, followed by Model 3 and so on. Consumer satisfaction indices for Models 2 and 4 are similar and they have the best overall reliability rating among all five models.

In this section, we first discuss in detail the results from the multiple vehicle CSI modeling of TS{1,2,3} and compare the CSI ranking against (11). Next we summarize the results for all possible model combinations and choose the best CSI model based on statistical testing.

7.1. CSI Model for TS{1,2,3}

The GA parameters for solving (10) for TS{1,2,3} are mostly the same as those specified for NSGA-II in Section 5 except the population size and number of generations which are set to 500 and 10000 respectively. Constraints are handled using the penalty-parameter-less approach described in [25]. The optimal CSI model obtained by solving Equation (10) for Models 1, 2, and 3 is

$$\begin{aligned} CSI_{\{1,2,3\}} &= T_1 + T_2 \times T_3 \times T_4 + T_5 + T_6 \\ &= X_1^{0.5226} X_5^{0.7215} + X_2^{0.8558} X_3^{0.6597} X_4^{0.0349} X_5^{0.8129} \\ &\quad + X_1^{0.7226} X_5^{0.5773} + X_1^{0.0255} X_3^{0.2429} X_5^{0.1124} \end{aligned} \tag{12}$$

The optimized CSI model $CSI_{\{1,2,3\}}$ is now used to evaluate the CSI values for all the customers in TS{1,2,3}. Figure 5 shows the normalized CSI distributions obtained for these vehicle models. The normalization is performed over all the customers in TS{1,2,3}. Normal distribution curves are fitted to the distributions in Figure 5 to clearly show the location of means. From Figure 5, it can be clearly seen that our primary goal of low variance and low-skewness distribution is met for all three vehicle models. The mean of each distribution can therefore be considered as overall CSI for each vehicle model. The inset in Figure 5 shows the relative positions of the means of the three distributions clearly revealing the satisfaction rating in the order,

$$CSI_3 \prec CSI_1 \prec CSI_2. \tag{13}$$

Here, mean of the CSI distribution of Model v has simply been denoted by CSI_v. The above agrees with the Consumer Reports rating in (11).

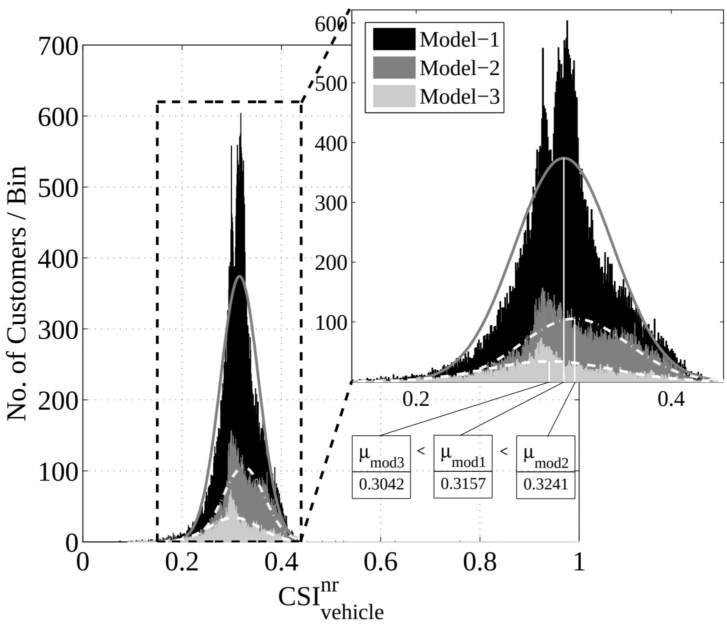


Figure 5: Normalized CSI distributions obtained using $CSI_{\{1,2,3\}}$ for Models 1, 2 and 3. The optimized objective function value is $|\mu_1 - \mu_2| + |\mu_1 - \mu_3| + |\mu_2 - \mu_3| = 0.039803$.

7.2. Statistical Hypothesis Testing

Table 5 shows some statistical characteristics of the three CSI distributions shown in Figure 5. Numerically, the mean CSI values for the three vehicle models are observed to be very close. This raises concerns regarding the statistical significance of (13). For analyzing this further, let us consider

Table 5: Statistical characteristics of the three CSI distributions shown in Figure 5.

Model No.	Mean μ_m	Std. dev. σ_m	Skewness $\lvert g_m \rvert$
Model 1	0.315774	0.038656	0.238562
Model 2	0.324148	0.041722	0.245239
Model 3	0.304247	0.049920	0.207301

the following null hypothesis,

$$H_0 : M_m - M_n = 0, \tag{14}$$

where M_m and M_n are the *population* means for Models m and n respectively. The use of Welch's t-test has been suggested in [26] for independent or unpaired (in statistical hypothesis testing terminology) samples with unequal sizes and variances. This test is basically an extension of the Student's two sample t-test. The t statistic is modified to accommodate unequal sample sizes as follows:

$$t = \frac{\mu_m - \mu_n}{\sqrt{\frac{\sigma_m^2}{c_m} + \frac{\sigma_n^2}{c_n}}}. \tag{15}$$

The degree of freedom (ν) expression is calculated using,

$$\nu = \frac{\left(\frac{\sigma_m^2}{c_m} + \frac{\sigma_n^2}{c_n}\right)^2}{\frac{\sigma_m^4}{c_m^2(c_m-1)} + \frac{\sigma_n^4}{c_n^2(c_n-1)}} \tag{16}$$

Note that μ and σ here are the *sample* mean and standard deviation respectively. As is customary in such tests, we use $\alpha = 5\%$ significance level. Table 6 shows the results of three Welch t-tests performed on all pairs of Models 1, 2 and 3. Since none of the confidence intervals encloses the hypothesized mean difference value of zero and the difference between sample means for all vehicle model pairs lie within the confidence interval, the null hypothesis H_0 can be rejected in all three tests with 95% confidence. The alternate hypothesis stating that the difference between the means is statistically significant for all three model pairs is hence accepted in all cases. The implication of each test is summarized in the last column of Table 6.

Table 6: Welch's t-test statistics.

Model Pair {m,n}	t	ν	95% confidence interval ($M_m - M_n$)	$\mu_m - \mu_n$	Implication
{1,2}	-28.72	66080.89	(-0.0089,-0.0078)	-0.008374	$CSI_1 \prec CSI_2$
{1,3}	15.35	5246.97	(0.0101, 0.0130)	0.011527	$CSI_1 \succ CSI_3$
{2,3}	26.00	5663.99	(0.0184, 0.0214)	0.019901	$CSI_2 \succ CSI_3$

7.3. CSI Models for Remaining Training Sets

In this section, we solve (10) for all training sets. Additionally, regardless of the number of vehicle models used, we apply the obtained CSI model in each case to all five vehicle models to test the generality of the obtained function[1]. Statistical tests are also performed in all cases. The procedure for obtaining the CSI precedence order and statistical testing remains exactly the same as in the above section.

Table 7 shows the overall CSI precedence relations obtained from all 26 possible training sets. The first column in the table shows the category of the TS, i.e. the number of vehicle models involved. Actual optimized values for the objective function in (10) are shown in the second to last column. The last column shows the scaled objective values obtained by dividing the actual objective value by the number of summation terms in the objective function. This scaled value is a measure of how well the corresponding CSI model differentiates between the vehicle model pairs.

It is to be noted that 24 out of 26 training sets yielded the correct CSI ordering against (11). From (11) it is seen that Models 2 and 4 have approximately similar ratings, hence the relative positions of CSI_2 and CSI_4 are ignored in the table when determining the correctness of the CSI precedence relations. Two training sets in II-TS category (highlighted in Table 7) yielded unsatisfactory results (incorrect CSI ordering or infeasible CSI model): an indication that two vehicle models as training sets may be not be sufficient for generalizing the obtained CSI model for other vehicle models. On the other hand, four or five training sets involve larger datasets and therefore high computational cost. Based on these simple observations we conclude that three training sets are sufficient for generalization in this study. Even within the III-TS category, we find that TS$\{2, 4, 5\}$ gives the best scaled objective value. Therefore, we shall use $CSI_{\{2,4,5\}}$ in the later sections to show some applications of the proposed CSI modeling technique. Its functional form is given by,

$$\begin{aligned} CSI_{\{2,4,5\}} &= T_1 \times T_2 \times T_3 \times T_4 \times T_5 \times T_6 \\ &= X_1^{0.0665} X_2^{1.9547} X_3^{5.9082} X_4^{0.5356} X_5^{2.6382} X_6^{5.9102}. \end{aligned} \tag{17}$$

[1]This is the reason for calling vehicle model combinations as training sets.

Table 7: Overall CSI precedence relations for all five vehicle models obtained from all possible vehicle model combinations. CSI models which are infeasible or which resulted in incorrect CSI precedence relations (against (11)) are marked in gray. The best CSI model is shown in **bold**.

Model Training Sets		Predicted CSI precedence (Training Sets + Test Sets)	Objective Value ($\times 10^{-2}$)	
			Actual	Scaled
II-TS	{1,2}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_4 \prec CSI_2$	1.3847	1.3847
	{1,3}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	1.3820	1.3820
	{1,4}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	1.3770	1.3770
	{1,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	3.2238	3.2238
	{2,3}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_4 \prec CSI_2$	2.0195	2.0195
	{2,4}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_4 \prec CSI_2$	0.9831	0.9831
	{2,5}	$CSI_5 \prec CSI_3 \prec CSI_2 \prec CSI_1 \prec CSI_4$	NA	NA
	{3,4}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	2.3566	2.3566
	{3,5}	No Feasible Solution Found	NA	NA
	{4,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	4.2273	4.2273
III-TS	{1,2,3}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_4 \prec CSI_2$	3.9803	1.3026
	{1,2,4}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_4 \prec CSI_2$	2.6595	0.8865
	{1,2,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	6.5799	2.1933
	{1,3,4}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	4.6536	1.5512
	{1,3,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_4 \prec CSI_2$	3.6520	1.2173
	{1,4,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	8.0151	2.6717
	{2,3,4}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	4.7065	1.5688
	{2,3,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_4 \prec CSI_2$	5.6200	1.8733
	{2,4,5}	$\mathbf{CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4}$	**8.4578**	**2.8192**
	{3,4,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	5.8007	1.9335
IV-TS	{1,2,3,4}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	7.4282	1.2380
	{1,2,3,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_4 \prec CSI_2$	9.1151	1.5191
	{1,2,4,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	12.3547	2.0591
	{1,3,4,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	9.2013	1.5335
	{2,3,4,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_2 \prec CSI_4$	10.5236	1.7539
V-TS	{1,2,3,4,5}	$CSI_5 \prec CSI_3 \prec CSI_1 \prec CSI_4 \prec CSI_2$	14.4659	1.4465

7.3.1. A Note on Hypothesis Testing

For the 24 training sets that yielded the correct CSI ordering, the Welch's t-test is performed on all vehicle model pairs. The null hypothesis is the same as in (14). However, regardless of the number of vehicle models used, this null hypothesis is tested on all possible model pairs ($^5C_2 = 10$). Thus, a total of 240 tests are performed. We are able to reject the null hypothesis in 239 cases. The failed test case corresponds to TS$\{2, 3\}$: another indication against using a training set with two vehicle models. The analysis is same as that described in Section 7.2 and hence is omitted here for brevity.

7.4. Sensitivity Analysis

An important part of CRM is to study what factors affect customer satisfaction to what extent. Understanding this gives crucial insights for attracting new customers and retaining old ones. In mathematical terms this study is referred to as sensitivity analysis. In this section we investigate the sensitivity of the CSI models obtained in the III-TS, IV-TS and V-TS categories of Table 7 to the transformed features X_i. Mathematically, the partial derivative

$$S_i = \frac{\partial f}{\partial X_i}$$

gives the *local* sensitivity with respect to the i-th transformed feature, where f the functional form of the CSI model under consideration.

In Section 1 we hypothesized that the satisfaction distribution is composed of (two or) three nearly normal distributions of (two or) three statistically differing populations - dissatisfied, satisfied and extremely satisfied customers. Here, we are specifically interested in how important these transformed features are for the extremely satisfied set of customers because ultimately they are responsible for the affective and behavioral components of satisfaction, which cannot be measured in quantitative terms. To extend the study over the entire market we choose the top 10% customers (by CSI value) of Model 2, one of the best vehicle models among the five considered in this paper, thus identifying the extremely satisfied customer set in the market. The sensitivities S_i are averaged over this customer set to get $\overline{S}_i$ for all 16 CSI models in III-TS, IV-TS and V-TS categories of Table 7. The average sensitivities are ordered according to their decreasing numeric values in Table 8 and classified into two groups of high sensitivity (H) and low sensitivity (L).

Table 8: Ordering of average CSI model sensitivities ($\overline{S}_i$) for top 10% customers (by CSI value) of Model 2. The high sensitivity (H) groups and low sensitivity (L) groups are also shown.

Training	CSI	Avg. Sensitivity Ordering ($\overline{S}_i$)		
Set (TS)	Model	Group H		Group L
{1,2,3}	$CSI_{\{1,2,3\}}$	$\overline{S}_3 > \overline{S}_5 > \overline{S}_1$	$>$	$\overline{S}_6 > \overline{S}_4 > \overline{S}_2$
{1,2,4}	$CSI_{\{1,2,4\}}$	$\overline{S}_5 > \overline{S}_5 > \overline{S}_2$	$>$	$\overline{S}_4 > \overline{S}_6 > \overline{S}_3$
{1,2,5}	$CSI_{\{1,2,5\}}$	$\overline{S}_3 > \overline{S}_1 > \overline{S}_5$	$>$	$\overline{S}_6 > \overline{S}_2 > \overline{S}_4$
{1,3,4}	$CSI_{\{1,3,4\}}$	$\overline{S}_3 > \overline{S}_6 > \overline{S}_5$	$>$	$\overline{S}_2 > \overline{S}_4 > \overline{S}_1$
{1,3,5}	$CSI_{\{1,3,5\}}$	$\overline{S}_5 > \overline{S}_1 > \overline{S}_2$	$>$	$\overline{S}_3 > \overline{S}_6 > \overline{S}_4$
{1,4,5}	$CSI_{\{1,4,5\}}$	$\overline{S}_6 > \overline{S}_3 > \overline{S}_5$	$>$	$\overline{S}_2 > \overline{S}_1 > \overline{S}_4$
{2,3,4}	$CSI_{\{2,3,4\}}$	$\overline{S}_6 > \overline{S}_3 > \overline{S}_2$	$>$	$\overline{S}_5 > \overline{S}_4 > \overline{S}_1$
{2,3,5}	$CSI_{\{2,3,5\}}$	$\overline{S}_2 > \overline{S}_5 > \overline{S}_3$	$>$	$\overline{S}_1 > \overline{S}_4 > \overline{S}_6$
{2,4,5}	$CSI_{\{2,4,5\}}$	$\overline{S}_3 > \overline{S}_6 > \overline{S}_5$	$>$	$\overline{S}_2 > \overline{S}_4 > \overline{S}_1$
{3,4,5}	$CSI_{\{3,4,5\}}$	$\overline{S}_6 > \overline{S}_2 > \overline{S}_3$	$>$	$\overline{S}_5 > \overline{S}_4 > \overline{S}_1$
{1,2,3,4}	$CSI_{\{1,2,3,4\}}$	$\overline{S}_6 > \overline{S}_3 > \overline{S}_5$	$>$	$\overline{S}_4 > \overline{S}_2 > \overline{S}_1$
{1,2,3,5}	$CSI_{\{1,2,3,5\}}$	$\overline{S}_1 > \overline{S}_5 > \overline{S}_3$	$>$	$\overline{S}_6 > \overline{S}_4 > \overline{S}_2$
{1,2,4,5}	$CSI_{\{1,2,4,5\}}$	$\overline{S}_6 > \overline{S}_3 > \overline{S}_5$	$>$	$\overline{S}_2 > \overline{S}_1 > \overline{S}_4$
{1,3,4,5}	$CSI_{\{1,3,4,5\}}$	$\overline{S}_2 > \overline{S}_5 > \overline{S}_6$	$>$	$\overline{S}_3 > \overline{S}_4 > \overline{S}_1$
{2,3,4,5}	$CSI_{\{2,3,4,5\}}$	$\overline{S}_2 > \overline{S}_6 > \overline{S}_5$	$>$	$\overline{S}_3 > \overline{S}_1 > \overline{S}_4$
{1,2,3,4,5}	$CSI_{\{1,2,3,4,5\}}$	$\overline{S}_6 > \overline{S}_5 > \overline{S}_4$	$>$	$\overline{S}_3 > \overline{S}_2 > \overline{S}_1$

The frequency of occurrence of the transformed features (with respect to which the average sensitivities are calculated) in groups H and L are shown in Figure 6. The following can be inferred from the figure:

(a) For the majority of extremely satisfied customers, the features X_3 (total cost), X_5 (average miles between visits) and X_6 (sum of severity ratings) are the most important factors governing their perception towards vehicle quality and reliability.

(b) The features X_1 (total number of visits) and X_2 (total waiting time of customer during vehicle service) are the next most important factors.

(c) The feature X_4 (average number of days between successive visits) is the least important factor for extremely satisfied customers.

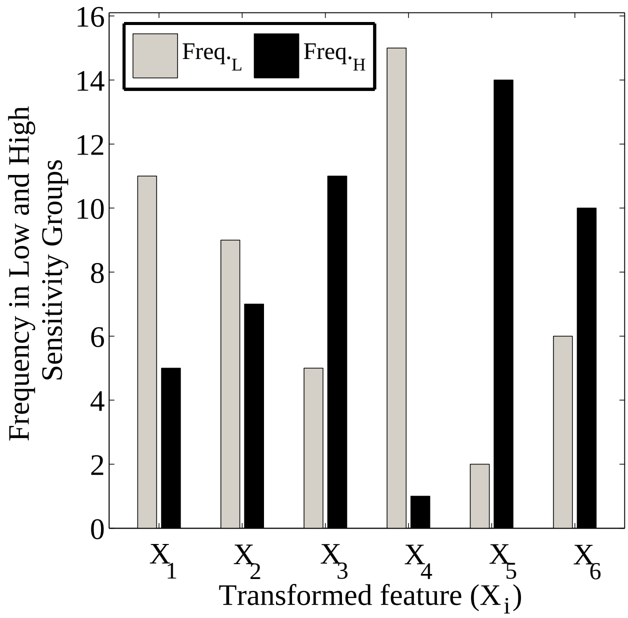


Figure 6: Bar graph showing the frequencies Freq_L and Freq_H with which the average sensitivity with respect to feature X_i occurs in L and H groups respectively.

Identification of relative importance of different features in the perception of a customer as outlined above remains as valuable information to designers. This information can be used by designers to prioritize design decisions. For example, improvements to design decisions which are directly related to features in the H-group of a vehicle model should be given preference over design decisions that effect features belonging to the L-group.

8. Applications of CSI Model

In this section we discuss two applications of the quantitative CSI models which are significant for a customer relations manager to take progressive decisions regarding design enhancements and future market strategy for all five vehicle models. For the reasons mentioned in Section 7.3 we use the CSI model given in (17) obtained using TS{2,4,5}.

8.1. Deriving Classification Rules

Using a quantitative CSI model and classification tree learning algorithms, we can identify a set of rules to clearly distinguish between any two sets of customers. Identification of such rules for classifying dissatisfied and satisfied customers is of prime importance for making future design enhancements to the product. More importantly, this technique allows automobile companies to make buyback offers to individuals in the dissatisfied set for ensuring customer retention and increasing customer loyalty. The procedure is illustrated here for Model 2. Consider Figure 7, which is a schematic representation of the CSI distribution for Model 2 obtained using $CSI_{\{2,4,5\}}$ in (17).

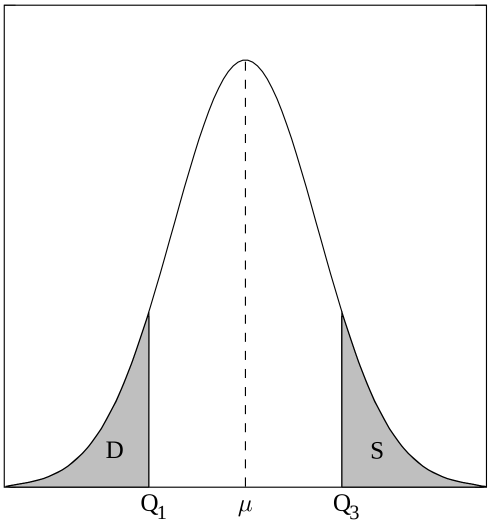


Figure 7: A schematic of the CSI distribution for Model 2 showing regions of satisfied (S) and dissatisfied (D) customers. Q_1 and Q_3 represent the first and the third quartiles respectively.

The customers are divided into four equal groups[2]. Customers with CSI rating below the first quartile Q_1 are labeled as Dissatisfied (Class D) and those with CSI rating above the third quartile Q_3 are labeled as Satisfied (Class S). MATLAB's `classregtree` algorithm is used to generate the classification tree in Figure 8. The features x_i extracted in Section 3 are used

[2]This is for illustration only. The actual choice depends on the purpose of the analysis.

as predictors for classifying customers as either belonging to D or S. The transformed features are not used here since their values are normalized and are of no relevance to a practitioner. At the root node of the classification tree we have equal number of S and D customers. The algorithm optimally branches a node such that at each sub-node the number of customers from one class (D or S) decreases and those from the other class increases. The process is continued until the number of customers from either of the two classes falls below 1%. This procedure is called pruning.

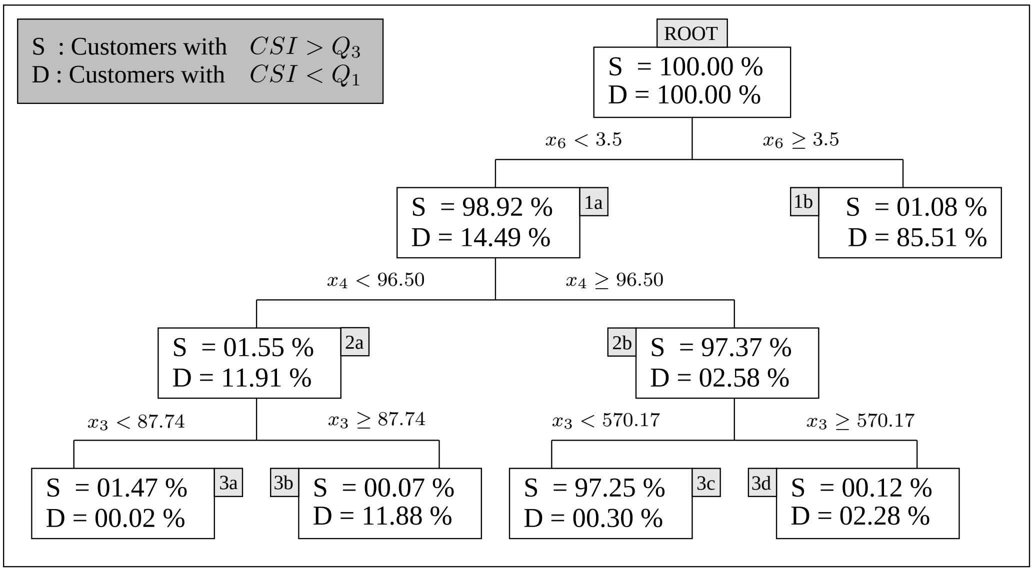


Figure 8: Classification tree for satisfied and dissatisfied customers of Model 2.

The classification tree reveals some very interesting rules for classifying satisfied and dissatisfied customers:

Rule-I: 97.25% customers from Class S follow the nodal path: `ROOT` → `1a` → `2b` → `3c`. This means almost all customers for whom,

$$[x_6 < 3.5 \texttt{ AND } x_4 \geq 96.5 \texttt{ AND } x_3 < 570.17] \text{ are satisfied.}$$

Rule-II: 85.51% customers from D follow the nodal path: `ROOT` → `1b`. Another 11.88% customers from D follow the path: `ROOT` → `1a` → `2a` → `3b`. Thus for 97.39% (85.51% + 11.88%) of dissatisfied customers,

$$[(x_6 \geq 3.5) \texttt{ OR } (x_6 < 3.5 \texttt{ AND } x_4 < 96.5 \texttt{ AND } x_3 \geq 87.74)].$$

Rule-I says that for the period considered in this study, if the following:

1. Sum of severity ratings of all repairs to the customer's vehicle is less than 3.5,
2. Average number of days between successive customer visits is more than 96.5 days, and
3. Total cost of repairs for the entire repair period is less than $570.17.

are *all* true for a customer, then it is very likely that he/she is satisfied with Model 2.

Similarly, Rule-II says that a customer is very likely to be dissatisfied with Model 2 if either,

1. Sum of severity ratings of all repairs to the customer's vehicle is more than 3.5,

or *all* of the following are true:

1. Sum of severity ratings of all repairs to the customer's vehicle is less than 3.5,
2. Average number of days between successive customer visits is less than 96.5 and
3. Total cost of repairs for entire service period is more than $87.74.

Clearly, such classification rules carry a lot of significance in data mining exercises relevant to CRM applications.

8.2. Identification of Critical Field Failures

A large customer base is also associated with a large number of repair types or failures. Table 1 shows the number of unique field failures in the five vehicle models considered. For planning future design enhancements, it becomes necessary to identify critical failures and deemphasize insignificant ones. Usually, this decision is based on the frequency or severity of the failures and on market feedback. In this section, we propose a method for prioritizing different field failures on the basis of the percentage improvement in overall satisfaction that can be achieved by eliminating (or reducing) their occurrence through design improvements. We call this quantity the CSI Improvement Potential (CIP) value of the field failure regardless of whether the corresponding field failure is completely or partially removed.

8.2.1. Procedure for Evaluating CIP

Let $\mathbf{CSI}_{orig}$ be the $C \times 1$ column vector of unnormalized CSI values obtained using any of the CSI models (say $CSI_{\{.\}}$) discussed in Section 7.3 for any given vehicle model with C customers. Then the procedure for calculating the CIP for 100% reduction of failures with repair code r_c is as follows:

Step 1: Remove all claims (for 100% reduction) containing repair code r_c from the service dataset. Some customers with only failures corresponding to r_c will be deleted altogether in the process. Let the number of such customers be Z. We will later compensate this reduction in the number of customers.

Step 2: Extract all six features x_i discussed in Section 3 to obtain the feature matrix FM of size $(C - Z) \times 6$. It is to be noted that reduction of any kind of failure for a customer should improve his/her satisfaction. However, when evaluating x_4 and x_5 for customers whose first or last visit has the repair code r_c, it is possible that the decrease in numerator of Equation (2) and/or Equation (3) is numerically greater than the decrease in denominator, thus decreasing the CSI value instead of increasing it. To avoid this, the numerator is not changed.

Step 3: Compensate for the deleted customers (who are now problem free) by adding Z instances of hypothetical customers with feature vector $[\min(x_1)\ \min(x_2)\ \min(x_3)\ \max(x_4)\ \max(x_5)\ \min(x_6)]$ to FM. This ensures that these customers are the most satisfied among the C customers.

Step 4: Normalize columns of x_i's in FM (now of size $C \times 6$) as described in Section 4 to get the transformed feature matrix FM^{tr}.

Step 5: Obtain $\mathbf{CSI}_{new}$, a $C \times 1$ column vector of unnormalized CSI values, using $CSI_{\{.\}}$ on FM^{tr}.

Step 6: Normalize the $2C \times 1$ sized column vector $\mathbf{CSI} = [\mathbf{CSI}_{orig}\ \mathbf{CSI}_{new}]^T$ between 0 and 1 using $\min(\mathbf{CSI})$ and $\max(\mathbf{CSI})$ to obtain $\mathbf{CSI}^{nr} = [\mathbf{CSI}^{nr}_{orig}\ \mathbf{CSI}^{nr}_{new}]^T$. This allows unbiased comparison between the original and new CSI distributions.

Step 7: The CIP for 100% reduction of failures with repair code r_c can now be calculated as,

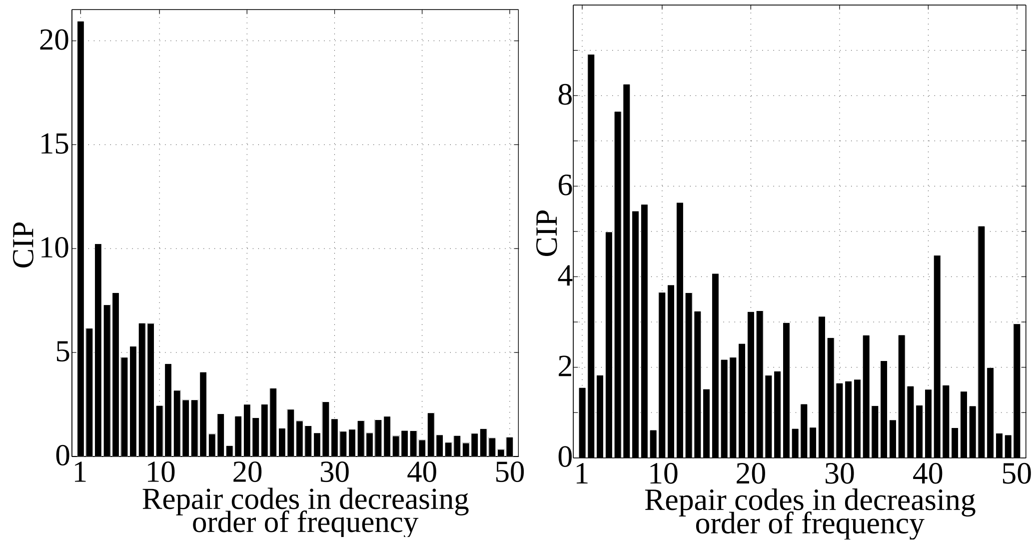
$$CIP = \frac{\text{mean}(\mathbf{CSI}^{nr}_{new}) - \text{mean}(\mathbf{CSI}^{nr}_{orig})}{\text{mean}(\mathbf{CSI}^{nr}_{orig})} \times 100\%. \tag{18}$$

8.2.2. Prioritizing Failures based on CIP

The ability to calculate CIP allows us to prioritize failures based on their impact on the CSI rating. Rectification of root causes of high priority failures leads to improvement in the design and more importantly improves the customer perception of the product. Traditionally, prioritization is done based on the frequency of occurrence of failures. Here, we investigate whether such an approach really improves the satisfaction proportionately. Models 1 and 5 are chosen for this study as they are the worst rated models in the non-luxury and luxury segments respectively (see Table 1 and Equation (11)). Figure 9 shows bar charts of CIP values corresponding to 100% removal of the top 50 most frequently occurring repair codes or failures (arranged in decreasing order of frequency along the X-axis) in Models 1 and 5. The CIP values are calculated using $CSI_{\{2,4,5\}}$. Note how the assumption that frequency of occurrence is directly proportional to the improvement in satisfaction holds good for Model 1 but not for Model 5. A viable approach for quantitatively identifying high-priority field failures, like the one presented here, is very useful for making progressive improvements to the vehicle design while ensuring that the customers are maximally satisfied with each design change.

8.2.3. CIP for Partial Reduction of Failures

Although the highest improvement in customer satisfaction is achieved by the complete removal of high-priority field failures, it may not always be practical. In such cases, it may be useful to study the effect of partial reduction of failures. Important insights into the nature of CIP can be gained in the process. Model 1 is used here for illustrating the CIP behaviour with respect to the most frequent repair code. Again the CSI model $CSI_{\{2,4,5\}}$ is used for all CIP calculations. Figure 9(a) reveals that the complete removal of the most frequent repair code leads to around 20% improvement in satisfaction. On the other hand, Figure 10 shows the progressive improvement in satisfaction for every 10% additional removal of the most frequent repair code, staring from 0% (where obviously $CIP = 0\%$) to 100% (where $CIP \approx 20\%$ as seen in Figure 9(a)). Figure 10 is generated using a sampling method. For



(a) (b)

Figure 9: CIP values corresponding to 100% removal of the top 50 most frequent repair codes for (a) Model 1 and (b) Model 5.

CIP calculation of any intermediate $p\%$ removal, a sample of 100 datasets is first created by randomly removing $p\%$ of the most frequent repair code. Thereafter, the CIP is calculated for all 100 datasets and a corresponding box-whisker ordinate is plotted at $p\%$ abscissa. The lower and upper edges of rectangular box represents the first and the third quartiles respectively of the 100 CIP values, the circle in the rectangular box represents their median value, and the whiskers represent the extreme CIP values.

8.2.4. Partial Reduction Plans

Figure 11 shows the CIP values for every 10% partial reduction of the top five most frequent repair codes (named r_1 through r_5) occurring in Model 1. The plots are obtained using exactly the same procedure as described above except that only the median values are shown for the box-whisker ordinates for clarity. It can be verified that the CIP values at 100% reduction of each repair code correspond to the first five bars in Figure 9(a). Note that despite using a highly non-linear CSI function given by $CSI_{\{2,4,5\}}$ from (17), the CIP increases approximately linearly with the amount of reduction for all five repair codes. More importantly, note that different repair codes give

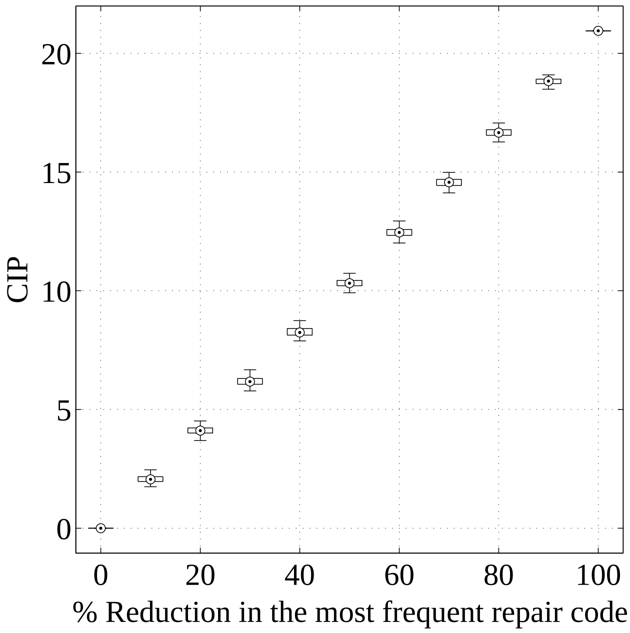


Figure 10: CIP values for every 10% partial reduction of the most frequent repair code in Model 1.

different rates of improvement in satisfaction.

With this information in hand, the next question to ask is what combination of repair codes should be used and how much reduction should be targeted for each repair code to achieve a desired level of improvement in customer satisfaction. Typically, the number of repair codes associated with each vehicle model, as shown in Table 1, is very high. For illustration we consider five hypothetical Partial Reduction Plans (PRPs) shown in Table 9. For example, PRP-I represents a 10% reduction in repair code r_1, 20% in r_2, 30% in r_3, 20% in r_4 and 10% reduction in r_5.

The sampling method is again used here for calculating the CIP value for each PRP shown in Table 9. Hundred datasets are created by randomly removing the specified percentages of all five repair codes. The corresponding CIP values are calculated as before. The median of these CIP values for all five hypothetical PRPs are shown in the first row of Table 10. We now estimate the CIP values for the same PRPs using Figure 11 by adding the individual CIP values of the five repair codes obtained by the amount of

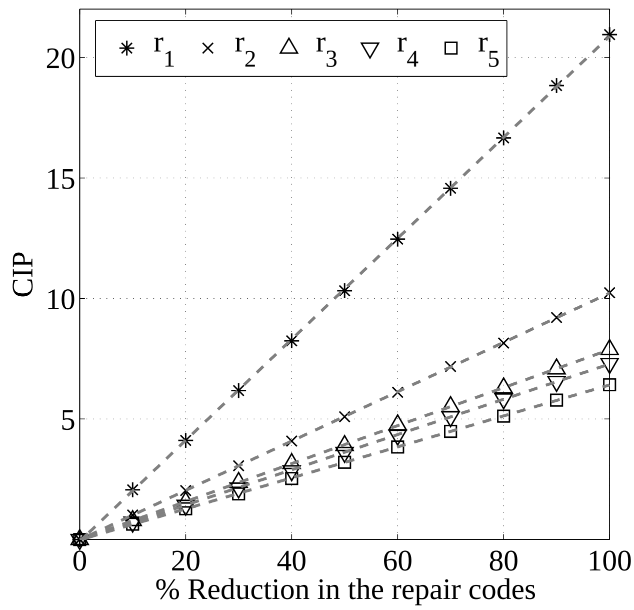


Figure 11: CIP median values for every 10% partial reduction of the top five most frequent repair codes in Model 1.

reduction specified by the PRP. For example, in case of PRP-I,

$$\begin{aligned} CIP_{PRP\text{-}I} &= CIP_{r_1}(10\%) + CIP_{r_2}(20\%) + CIP_{r_3}(30\%) \\ &\quad + CIP_{r_4}(20\%) + CIP_{r_5}(10\%) \\ &= 2.059\% + 2.034\% + 2.358\% + 1.426\% + 0.635\% \\ &= 8.512\%. \end{aligned} \tag{19}$$

This is known as the principle of superposition. The second row of Table 10 shows that this principle approximately holds for CIP calculations with different reduction plans. This is due to the fact that most failure types are independent of each other. The small variation is seen due to the finite number (100) of samples used for CIP calculation.

9. Conclusion

In this study, we have demonstrated how starting with vehicle service and sales data, a plethora of valuable information can be derived about perceived customer satisfaction indicators. First, a single-objective optimization

Table 9: Five hypothetical Partial Reduction Plans (PRPs) using the top five most frequent repair codes of Model 1.

	PRP-I	PRP-II	PRP-III	PRP-IV	PRP-V
r_1	10% ↓	10% ↓	10% ↓	10% ↓	10% ↓
r_2	20% ↓	10% ↓	10% ↓	10% ↓	20% ↓
r_3	30% ↓	20% ↓	10% ↓	10% ↓	10% ↓
r_4	20% ↓	20% ↓	20% ↓	10% ↓	10% ↓
r_5	10% ↓	20% ↓	20% ↓	20% ↓	10% ↓

Table 10: Comparison between median CIP values obtained by the sampling method and by the principle of superposition using Figure 11.

	PRP-I	PRP-II	PRP-III	PRP-IV	PRP-V
CIP Median (using 100 samples)	8.58	7.39	6.59	5.87	6.25
CIP Median (from Figure 11)	8.51	7.32	6.55	5.85	6.23

procedure has been suggested to develop Customer Satisfaction Index (CSI) models for different vehicle models as a function of six important features extracted from the service data. This customer level CSI modeling approach gave us an understanding of the objective space and difficulties associated. Consequently, we developed the multiple vehicle CSI modeling technique through which a single CSI function can allow us to rank different vehicle models in increasing order of customer satisfaction. The ranking of five different vehicle models obtained by our procedure has been statistically validated to be significant and also verified against the ranking obtained from published Consumer Reports reliability ratings on the same set of vehicle models.

The generalizing capability of the proposed optimization approach has been tested by using a subset of vehicle models for the optimization task and testing the remaining vehicle models using the obtained CSI function. For the five vehicle models, the use of a maximum of three during optimization was found to be adequate to correctly rank all five vehicle models. This makes the use of the proposed procedure practically viable for handling a large number of vehicle models.

Once a CSI function is developed and verified, it can be used to rank the features that most significantly affect customer perception. For a particular vehicle model, we have identified the importance of different features with respect to the most satisfied customers. It has been observed that total cost of repair, average miles elapsed between two consecutive visits and the sum of severity ratings of repairs are the most important features causing the perceived satisfaction of a customer. Contrary to general belief, average number of days between successive visits has been found to be the least important factor in influencing customer satisfaction.

By identifying the extremely satisfied customers and dissatisfied customers from the obtained CSI function for a vehicle model, classification rules relating to the features that are responsible for influencing the customers have been obtained by using a classification tree algorithm. These rules can help the company in easily identifying dissatisfied customers for making buyback proposals.

Automobile industries are also interested in identifying field failures that are critically responsible for causing customer dissatisfaction. If such critical field failures can be identified, an effort to reduce them through future design improvements would be an indirect benefit of the CSI study. With the developed CSI function model, a number of field failures can be considered for this purpose. First, our analysis has identified a set of critical field failures that most significantly affect the CSI function and then we have demonstrated how partial or complete removal of such critical field failures can lead to improvement in perceived customer satisfaction.

All the above analysis, originating from the proposed CSI modeling approach, has been shown to provide useful insights for improving perceived customer satisfaction. But in the long run such information should also indirectly help in designing a better product.

Acknowledgements

The financial support and vehicle related data provided by India Science Lab, General Motors R&D are greatly appreciated. Authors thank Dr. Prakash G. Bharati, Dr. Pulak Bandyopadhyay and Dr. Pattada A. Kallappa for helpful discussions.

References

[1] ISO 9001:2008: Quality management systems - Requirements, ISO/TC 176 (2008).

[2] C. Cianfrani, J. Tsiakals, J. West, ISO 9001: 2008 Explained, ASQ Quality Press, 2009.

[3] L. Dube, K. Menon, Multiple roles of consumption emotions in post-purchase satisfaction with extended service transactions, International Journal of Service Industry Management 11 (3) (2000) 287–304.

[4] L. Berry, Marketing services: Competing through quality, Free Press, 2004.

[5] S. Kessler, Customer satisfaction toolkit for ISO 9001: 2000, American Society for Quality, 2002.

[6] R. Likert, A technique for the measurement of attitudes., Archives of psychology.

[7] S. Eroglu, K. Machleit, An empirical study of retail crowding: Antecedents and consequences, Journal of Retailing 66 (2) (1990) 201–221.

[8] R. Westbrook, A rating scale for measuring product/service satisfaction, The Journal of Marketing (1980) 68–72.

[9] The American Customer Satisfaction Index, ACSI, www.theacsi.org (2010).

[10] JDPower.com, J.D. Power and Associates, www.jdpower.com (2010).

[11] ConsumerReports.org, Consumers Union of U.S., Inc., www.consumerreports.org (2010).

[12] J. Robinson, S. Chukova, Estimating mean cumulative functions from truncated automotive warranty data, in: Communications of the Fourth International Conference on Mathematical Methods in Reliability, Methodology and Practice, Santa Fe, New Mexico, USA, 2004, pp. CD–ROM (4 pages).

[13] A. Parasuraman, V. Zeithaml, L. Berry, A conceptual model of service quality and its implications for future research, The Journal of Marketing (1985) 41–50.

[14] A. Parasuraman, V. Zeithaml, L. Berry, SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality, Retailing: Crit Concepts Bk2 64 (1) (2004) 140.

[15] S. Wu, Warranty data analysis: A review, Quality and Reliability Engineering International.

[16] S. Bandaru, K. Deb, V. Khare, R. Chougule, Quantitative modeling of customer perception from service data using evolutionary optimization, in: Proceedings of the 13th annual conference on Genetic and evolutionary computation, ACM, 2011, pp. 1763–1770.

[17] J.D. Power/What Car? 2009 UK Vehicle Ownership Satisfaction Study, www.testdriven.co.uk (2009).

[18] N. Anisor, D. Adela-Eliza, C. Luciana, The method used for measuring the customers' satisfaction, in: Proc. 9th WSEAS International Conference on Education And Educational Technology (EDU'10), 2010, pp. 197–200.

[19] J. Koza, Genetic programming: On the programming of computers by means of natural selection, MIT press, 1992.

[20] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multi-objective genetic algorithm: NSGA-II, IEEE Transactions on Evolutionary Computation 6 (2) (2002) 182–197.

[21] K. Deb, R. Agrawal, Simulated binary crossover for continuous search space, Complex Systems 9 (2) (1995) 115–148.

[22] K. Deb, Multi-objective optimization using evolutionary algorithms, Wiley, Chichester, UK, 2001.

[23] K. Deb, S. Gupta, Understanding knee points in bicriteria problems and their implications as preferred solution principles, Engineering optimization.

[24] ConsumerReports, ConsumersReports.org (April 2010).

[25] K. Deb, An efficient constraint handling method for genetic algorithms, Computer Methods in Applied Mechanics and Engineering 186 (2–4) (2000) 311–338.

[26] R. Lomax, An introduction to statistical concepts for education and behavioral sciences, Lawrence Erlbaum, 2001.