TRANSPORT
ISSN 1648-4142 print / ISSN 1648-3480 online
2014 Volume 29(1): 75–83
doi:10.3846/16484142.2014.898216

KEY FACTORS AFFECTING RAIL SERVICE QUALITY IN THE NORTHERN ITALY: A DECISION TREE APPROACH

Rocío De Oña[1], Laura Eboli[2], Gabriella Mazzulla[2]

[1]*Dept of Civil Engineering, University of Granada, Spain*
[2]*Dept of Civil Engineering, University of Calabria, Italy*

Submitted 6 April 2013; accepted 17 July 2013

Abstract. This work concerns with the analysis of transit service quality on the basis of the perceptions directly expressed by the passengers of the services. The transit services supporting the research are offered by rail operators of the Northern Italy, and particularly by regional and suburban lines connecting different towns of the hinterland of the city of Milan, and express lines connecting Milan with the Malpensa airport. The experimental data were collected in a survey conducted in May 2012, and addressed to a sample of more than 16,000 passengers. Passengers expressed their opinions about service characteristics such as safety, cleanliness, comfort, information, personnel. The tool chosen for evaluating service quality is a Classification and Regression Tree Approach (CART), useful for identifying the characteristics mostly influencing the overall service quality. We found that service characteristics like 'Windows and Doors Working', 'Courtesy and Competence on Board', 'Information at Stations', 'Punctuality of Runs', 'Courtesy and Competence in Station' and 'Regularity of Runs' mainly influence service quality.

Keywords: railway transit services, service quality attributes, service quality analysis, passengers' perceptions, CART methodology.

Reference to this paper should be made as follows: De Oña, R.; Eboli, L.; Mazzulla, G. 2014. Key factors affecting rail service quality in the Northern Italy: a decision tree approach, *Transport* 29(1): 75–83.
http://dx.doi.org/10.3846/16484142.2014.898216

Introduction

Assuring the quality of the public transportation system is an essential task for the transportation engineers and authorities. Transit service providers need to measure the performances of their service to verify the efficiency, effectiveness, and quality of the service. These measures can be used for monitoring the service, evaluating economic performance, administering the organization, developing service design standards and noting community benefits (TRB 2003). Especially evaluation of service quality is still a key issue in transport engineering, as transit is often argued as a promising travel mode to reduce dependence on automobile, thereby lessening traffic congestion, alleviating environmental pollution, and so on (Nocera 2011). Transportation contributes 26% to the overall CO_2 emissions, thus constituting the second biggest polluter in the EU (Nocera, Cavallaro 2011, 2012).

In the context of service quality assessment, the importance of 'user perception' has been highlighted and prioritized by numerous researchers. Service quality is defined as 'an antecedent of customer satisfaction' which in turn is based on 'user perception' (Wang, Shieh 2006).

In this paper, we propose an analysis of Service Quality (SQ) conducted on the basis of users' perceptions, expressed in terms of satisfaction and importance assigned to the various service characteristics. More specifically, the proposed work aims to an investigation of the influence of several service characteristics on the overall quality of a railway service. Particularly, we propose the use of a Classification and Regression Tree Approach (CART) to achieve these aims, retaining it as suitable for identifying the characteristics mostly influencing the overall service quality, and then the most convenient investments for improving the service.

After this brief introduction, we propose a concise literature review of works concerning methodologies for analysing customer satisfaction in public transport, in order to introduce the CART methodology as an alternative way for analysing transit service quality; CART is

Corresponding author: Gabriella Mazzulla
E-mail: *gabriella.mazzulla@unical.it*

Taylor & Francis
Taylor & Francis Group

described in the third section. The experimental context is presented in the fourth section; more specifically, we briefly describe the survey conducted for collecting the data, the sample of interviewed users, and the analysis of the importance and satisfaction rates expressed by the users. The fifth section is about the results obtained from the application of the CART methodology. Finally we propose a brief general discussion of the work.

1. Literature Review

Transit service quality can be measured by different approaches. The most common approach is based on transit users' opinions about the used services, where the different aspects of the service are rated by the users during a survey. An alternative approach entails using different variables of the transit system demand and operation to calculate the 'efficiency' indicators (e.g. Badami, Haider 2007; Lao, Liu 2009; Eboli, Mazzulla 2012a). There are also some studies integrating the two mentioned approaches (e.g. Sheth *et al.* 2007; Abreha 2007; Nathanail 2008; Eboli, Mazzulla 2011).

Concerning the approach based on transit users' opinions, we have to specify that users' judgments have been generally expressed in terms of perceptions of the used service (expressed in terms of satisfaction rates) and/or expectations about the service (expressed in terms of importance rates). Knowing the importance given by the users to each service characteristics is fundamental for understanding and defining the strategies for improving the quality of a service. Particularly, importance can be directly derived asking customers to rate each attribute on an importance scale, or it can be calculated by methods statistically testing the strength of the relationship of individual attributes with overall satisfaction (De Oña *et al.* 2013). Recent years have seen the development of these methods based on: traditional customer satisfaction surveys, in which users express their opinions by rating the various service characteristics (Cavana *et al.* 2007; Dell'Olio *et al.* 2010; Eboli, Mazzulla 2010, 2012b; Jen *et al.* 2011; Joewono, Kubota 2007; Habib *et al.* 2011; Pakdil, Aydın 2007; Weinstein 2000); stated preference surveys, in which the importance given by the users to the service attributes is indirectly derived by means of exercises based on the stated preferences according to which users express their opinions by choosing (or rating of ranking) alternative hypothetical services (Cirillo *et al.* 2011; Eboli, Mazzulla 2008a, b, 2010; Hensher, Prioni 2002; Hensher *et al.* 2003; Dell'Olio *et al.* 2011).

However, most of these models have their own model assumptions and pre-defined underlying relationships between dependent and independent variables. If these assumptions are violated, the model could lead to erroneous estimations of the likelihood of quality of service (De Oña *et al.* 2012). In this context, the CART can solve these inconveniences, being a non-parametric model with no pre-defined underlying relationship between the target (dependent) variable and the predictors (independent variables). In the field of transportation, CART was applied for the road safety analysis (Abdel-Aty *et al.* 2005; Council, Stewart 1996; Chang, Chen 2005; Chang, Wang 2006; Chen, Jovanis 2000; Kuhnert *et al.* 2000; Magazzù *et al.* 2006; Pande *et al.* 2010; Qin, Han 2008; Sohn, Shin 2001; Yan, Radwan 2006; Yan *et al.* 2010). The method has also been used for analysing other aspects of traffic engineering: some examples can be found in Washington, Wolf (1997), Washington *et al.* (1997), Hallmark *et al.* (2002). This approach was also adopted for analysing quality of service in industries different from transportation (e.g. Huang, Hsueh 2010), while Wong, Chung (2007) applied the methodology to the sector of air transportation. Concerning the analysis of transit service quality, this methodology was applied only by De Oña *et al.* (2012), who adopted the CART for identifying the key factors affecting the quality level of a bus service operating in Granada. Starting from this study, the authors propose the application of the CART to a railway service operating in a regional context of the Northern Italy.

2. Methodology

Decision trees are a data mining technique used to classify and predict a class variable. When the value of the target variable is discrete, a classification tree is developed, whereas a regression tree is developed for a continuous target variable. The CART method can be used for both the target variables. In this study, the target variable is discrete (rail SQ), and thus, a classification tree is used.

The development of a CART model begins with all the data concentrated in the root node, which is the node located at the top of the tree. This root node is divided into two child nodes on the basis of an independent variable (splitter) that maximizes the 'purity' of the child nodes. Then, each child node is recursively split until all of them are pure (all the cases are of the same class) or their 'purity' cannot be increased. The most famous splitting index is the Gini Index (Gini 1912), which measures the impurity of the node.

The impurity measure $I(t)$ at a node t may be defined as follows (SAS Institute Inc. 2004):

$$I(t) = 1 - \sum_{i=1}^{i} \left(\frac{n_i}{n} \right)^2, \qquad (1)$$

where: i is the number of classes in the target variable; n_i is the number of cases belonging to the class i; n is the total number of cases.

If a node is 'pure', all the observations in the node belong to one class, and the Gini Index or Impurity (node) will be equal to zero.

Then, we can define the split criterion based on the Gini Index as the Gini Reduction Criterion, which measures the 'worth' of each split in terms of its contribution toward maximizing the homogeneity through the resulting split. A set of candidate split rules are evaluated and ranked during the tree growth. If a split results in splitting of one parent node into B branches (in this case two branches, because CART model generates binaries trees), the 'worth' of that split may be measured as follows (SAS Institute Inc. 2004):

$$Worth = I(P) - \sum_{b=1}^{B} P(b) \cdot I(b), \quad (2)$$

where: $I(P)$ denotes impurity of the parent node; $P(b)$ denotes the proportion of observations in the node assigned to a branch b; $I(b)$ denotes the impurity of the node b.

Following this procedure, the maximal tree overfitting the data is created. To decrease its complexity and create simpler trees, pruning is realized according to a cost-complexity algorithm (Breiman *et al.* 1984) based on removing the branches adding little to the predictive value of the tree. After pruning a branch, if the increase in the misclassification cost is sufficiently lower than the decrease in the complexity cost, the branch will be pruned, and a new tree will be created.

The last step is to select an optimal tree from the pruned trees. Using the misclassification cost on the testing dataset (or an independent dataset), the optimal tree is the one having the least misclassification cost.

One of the methods for developing the model randomly divides the sample used in the training phase into k sets (k-fold cross validation). Sequentially, each subset is kept to be used as a testing set against the tree model generated by the remaining k–1 subsets. Thus, different k models are obtained, in which the accuracy of the classifications in the training set k–1 and in the testing subsets k can be evaluated and the optimal tree can be selected.

A more detailed description of the CART analysis and its applications can be found in Breiman *et al.* (1984).

One of the main advantages of the decision trees as opposed to other modelling methods is that they are presented as easily understandable visual branching images that provide effective *If–Then* rules. Every leaf of the decision tree corresponds to a decision rule that extracts very useful information about the data. It is a logic conditional structure starting in the root node with *If*, continuous with every variable that takes part in the tree growing making an *If* of the rule, and ends in the child nodes with *Then*, in which is associated the class of the target variable showing the highest number of cases in the analyzed child node.

Another valuable outcome provided by CART analysis is the value of the standardized importance of independent variables, which reflects the impact of such predictor variables on the model.

3. Experimental Context

We investigated on a railway service operating in the North of Italy, and specifically in the city of Milan. The analysed service offers different types of connections: 32 regional lines and 9 suburban lines connecting towns of the hinterland of Milan, and 2 express lines connecting Milan with the Malpensa airport. If we consider all the 43 lines, we have about 570000 passengers per day; more specifically, a regional line is used by about 25000/30000 passengers per day, while a line to Malpensa airport is used by about 8000 passengers per day. The maximum length of a line is about 150 km; this is the length characterizing the lines connecting the city of Milan and the hinterland (Lombardy region) to regions such as Piedmont and Veneto. The analysed lines offer a number from 35 to 83 runs per day, giving a service frequency of 2÷4 runs per hour.

A face-to-face survey was addressed to a sample of 16647 users (sample rate of about 3%), and realized in the month of May 2012. The interviews were conducted on board during the whole week (weekday, before a holyday, and holyday days) in a time slot between 6:00 AM and 10:00 PM.

Users answered a questionnaire structured into two main sections. The first section was about: general information (e.g. time period of the interview, train, line, station, and operator); socio-economic characteristics (e.g. gender, age, qualification, professional condition, and income); travel habits (e.g. trip scope and frequency, and ticket). The second section was specific about passengers' perceptions of the used services; users expressed importance and satisfaction rates, on a cardinal scale from 1 to 10, about 27 service quality factors concerning safety, cleanliness, comfort, service, information, personnel, etc.

About 52% of the sample was interviewed on trains of regional lines, 44% suburban lines, and the remaining 4% the Malpensa express services. The most part of users (85.5%) were interviewed in a weekday, 7.8% in a before a holiday day, and 6.6% in a holiday. About one third of the passengers were interviewed in the off-peak hours (31.2%), 26.5% in the afternoon peak-hours, 23.8% in the evening peak hours, and 18.5% in the morning peak-hours. 27.5% of the interviewed people travel by train for working, 28.3% for studying, and the remaining 44.2% travel for other purposes.

The sample is made up more of females. Most of the passengers are aged between 16 and 25, and another fair chunk is represented by people aged between 26 and 40. The major part of the sampled people are students, but a considerable part is composed of employees. More than half of the sample obtained a diploma of a secondary school of second level, and almost one third has a degree. About one fifth of the sample doesn't give any kind of information about income, while 40.0% has not a fixed income; people stating their income mainly belongs to a class between 1001 and 1500 EUR.

Passengers travel by train mainly for reaching the place of work or study (73.0%). More than half of the sample travels by train every day, but about 26% of passengers travel occasionally. People mainly purchase a travel card (66.0%), but about 30% travel using a one-way ticket (Table 1). Table 2 shows the average importance and satisfaction rates calculated from the rates expressed by the users about the 27 service attributes.

According to the users all the attributes are very important (having an average rate of importance around 8 and 9); only the attribute linked to the parking is considered relatively as less important (average importance rate of 7.3). The attributes considered as the most important are the three attributes concerning travel safety. On the other hand, the average satisfaction rates suggest that people are not very satisfied with the service, in fact only eight attributes have an average rate higher than the sufficiency (>6). The service characteristics considered as the most satisfying regard safety and personnel, while the characteristics judged as the less satisfying concern cleanliness. By observing the satisfaction rate about the overall service, we could state that users consider the quality of the service on the whole as almost sufficient.

Table 1. Sample characteristics

Characteristics	Statistics	
Gender	male	45.5%
	female	54.5%
Age	16÷25	43.1%
	26÷40	31.4%
	1÷65	21.6%
	>65	3.9%
Professional Condition	employee	(35.2%)
	manager	2.0%
	entrepreneur	1.3%
	freelancer	5.2%
	self-employed worker	4.3%
	unemployed	3.8%
	student	41.4%
	housewife	2.0%
	pensioner	4.1%
	other	0.7%
Income Level (EUR)	≤1000	8.9%
	1001÷1500	15.0%
	1501÷2000	7.1%
	2001÷3000	4.8%
	3001÷4000	1.7%
	>4000	1.5%
	no fixed income	40.0%
	no answer	21.0%
Qualification	degree	31.5%
	diploma of secondary school of second level	55.5%
	diploma of secondary school of first level	11.8%
	diploma of primary school	1.2%
Scope of Journey	work	36.8%
	studying	36.2%
	bureaucratic activities	3.0%
	personal activities	20.0%
	tourism	4.0%
Frequency of Journey	daily	57.8%
	weekly	15.9%
	occasionally	26.2%
Ticket Kind	one-way ticket	29.4%
	carnet	4.5%
	travel card	66.0%

Table 2. Importance and Satisfaction rates

Service aspect	Service quality attribute	Importance rate	Satisfaction rate
Safety	Travel Safety	9.17	7.41
	Personal Security on Board	9.20	6.60
	Personal Security at Station	9.20	6.32
Cleanliness	Cleanliness of Vehicles	8.71	5.09
	Cleanliness of Seats	8.77	4.99
	Cleanliness of Toilet Facilities	8.72	4.24
	Cleanliness of Stations	8.26	5.36
	Maintenance of Stations	8.10	5.36
Comfort	Crowding on Board	8.25	5.32
	Air-conditioning on Board	8.37	5.27
	Windows and Doors Working	8.27	5.66
Service	Fare/Service Ratio	8.79	4.98
	Frequency of Runs	8.75	5.92
	Punctuality of Runs	9.00	5.58
	Regularity of Runs	8.87	5.89
	Price Integration with PT	8.29	5.92
	Localization of Stations	8.27	6.62
Other	Parking	7.90	5.50
	Bicycle Transport on Board	7.28	5.81
	Facilities for Disabled	8.91	5.09
Information	Information at Stations	8.47	5.65
	Information on Board	8.40	5.41
	Complaints	8.25	5.13
	Info Connections with PT	8.14	5.22
Personnel	Courtesy and Competence on Board	8.27	6.57
	Ticket Inspection	8.06	6.22
	Courtesy and Competence in Station	8.28	6.32
	Overall service		5.70

4. Results

A decision tree was built to classify the overall service quality of the railway service of Milan, and to identify the variables playing a key role in the classification of this variable. For this purpose, the 27 attributes describing the service, shown in Table 2 with the respective importance and satisfaction rates, were used as independent variables of the model, and named as ITEM1 to ITEM27. To find out more applicable decision rules, the target variable (overall SQ) and the independent variables were re-coded in a reduced semantic scale, and specifically a three semantic scale comprising the rates from 1 to 4 as POOR, from 5 to 7 as FAIR, and from 8 to 10 as GOOD.

The CART used a 10-fold cross-validation of the sample, which gave us a precision ratio of the categorization of the variable class of 77.43%. This value is acceptable and higher than the values obtained in other studies in which decision trees were applied with similar objectives. As an example, Wong and Chung (2007) obtained 61.2% of accuracy, and de Oña *et al.* (2012) obtained values of accuracy of 59.7% and 62.2%.

This model produced 7 levels, 29 nodes and 15 terminal nodes (Figure), corresponding the latest to decision rules for identifying the overall SQ. The root variable generating the tree is ITEM15 'Regularity of Runs', which splits into two branches (*Node* 1 and *Node* 2). This splitter is the variable obtaining the maximum 'purity' of the two child nodes. Then, passengers having a POOR satisfaction with the 'Regularity of Runs' are on the left branch of the tree, while passengers having a FAIR or GOOD satisfaction with this variable are on the right branch of the tree.

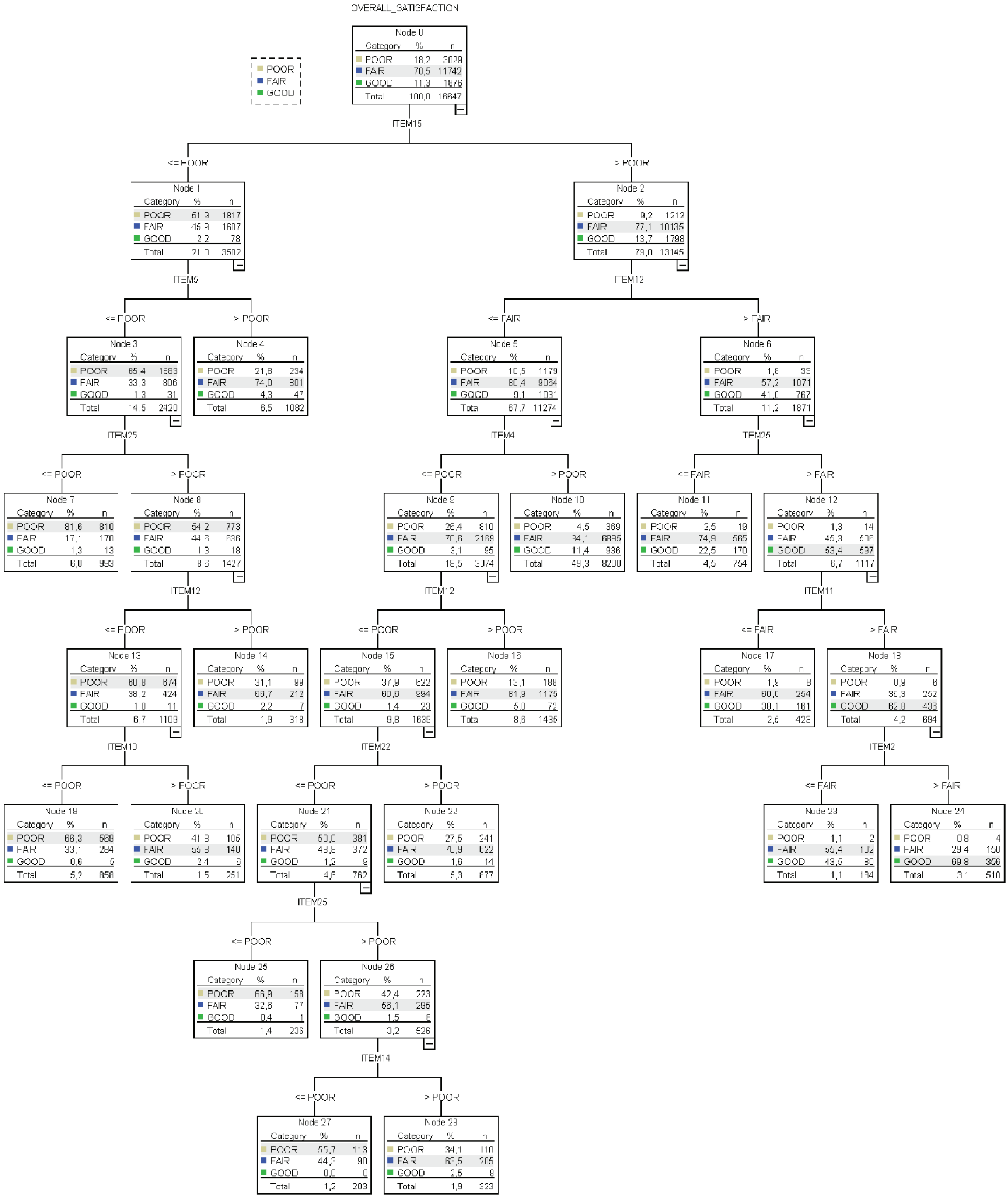

Figure. CART for overall satisfaction

Left branch gives 5 terminal nodes (4, 7, 14, 19 and 20). All these terminal nodes predict that passengers will rate the overall satisfaction as POOR or FAIR. This implies that a passenger, whose rate of satisfaction to the 'Regularity of Runs' is POOR, will never have an overall satisfaction with the service as GOOD, even if other variables are perceived with high satisfaction.

In turn, in *Node* 1 the tree splits by the ITEM5 'Cleanliness of Seats': if it is stated as FAIR or GOOD, the overall satisfaction is FAIR, with a probability of 74% (*Node* 4); in other cases the tree continuously grows according to the ITEM25 'Courtesy and Competence on Board'. If ITEM25 is rated as POOR, the overall satisfaction of the user will be POOR (terminal *Node* 7, 81.6%). Also in this left branch of the tree ITEM12 'Fare/Service Ratio' and ITEM10 'Air-conditioning on Board' were used as splitters for classifying the variable class.

The rest of the terminal nodes, found in the right branch of the tree (*Nodes* 10, 11, 16, 17, 22, 23, 24, 25, 27, and 28), are attributable to passengers with FAIR or GOOD satisfaction with the ITEM 15 'Regularity of Runs'. *Node* 2 is splitted according to the ITEM12 'Fare/ Service Ratio'. After *Node* 5 (passengers who consider ITEM 12 to be POOR or FAIR) six terminal *Nodes* are obtained (10, 16, 22, 25, 27, and 28), in which the overall satisfaction with the service is FAIR or lower. This implies that although the passengers are satisfied with ITEM15 (stated as FAIR or GOOD), if they have the impression that ITEM12 is POOR or FAIR, their global satisfaction with the service will not be GOOD.

On the other hand, when ITEM12 is considered to be GOOD (*Node* 6), four terminal *Nodes* are obtained (11, 17, 23 and 24) in which the overall satisfaction with the service is FAIR or higher. Moreover when also ITEM25 'Courtesy and Competence on Board', ITEM11 'Windows and Doors Working', and ITEM2 'Personal Security on Board' are stated with GOOD satisfaction, the overall satisfaction with the railway service will be GOOD at terminal *Node* 24 (69.8%).

Table 3 shows the 15 decision rules extracted by the decision tree building. The variables used to identify these rules are: ITEM15 'Regularity of Runs', ITEM5 'Cleanliness of Seats', ITEM25 'Courtesy and Competence on Board', ITEM12 'Fare/Service Ratio', ITEM10 'Air-conditioning on Board', ITEM4 'Cleanliness of Vehicles', ITEM22 'Clear and fast info on board', ITEM14 'Train punctuality', ITEM11 'Windows and Doors Working', and ITEM2 'Personal Security on Board'. Only one rule was found to imply a high probability that the overall satisfaction with the service will be GOOD (*Node* 24). If ITEM15 is perceived as FAIR or GOOD, and ITEM12, ITEM25, ITEM11 and ITEM2 as GOOD, the overall satisfaction with the service is likely to be GOOD (69.8%). On the contrary, four rules for POOR satisfactions and ten rules for FAIR satisfactions were identified. It can be observed that a POOR satisfaction with ITEM15 'Regularity of Runs' is the main cause for a POOR 'Overall Quality of Service', and also when the ITEM5 'Cleanliness of Seats' has a POOR satisfaction, this probability

Table 3. Rules for overall SQ

Node	Rule		Accuracy rate (%)
	IF...	THEN	
7	IF (ITEM15=POOR) AND (ITEM5=POOR) AND (ITEM25=POOR)	POOR	81.6
19	IF (ITEM15=POOR) AND (ITEM5=POOR) AND (ITEM25>POOR) AND (ITEM12=POOR) AND (ITEM10=POOR)	POOR	66.3
25	IF (ITEM15>POOR) AND (ITEM12<=FAIR) AND (ITEM4=POOR) AND (ITEM12=POOR) AND (ITEM22=POOR) AND (ITEM25=POOR)	POOR	66.9
27	IF (ITEM15>POOR) AND (ITEM12<=FAIR) AND (ITEM4=POOR) AND (ITEM12=POOR) AND (ITEM22=POOR) AND (ITEM25>POOR) AND (ITEM14=POOR)	POOR	55.7
4	IF (ITEM15=POOR) AND (ITEM5>POOR)	FAIR	74.0
14	IF (ITEM15=POOR) AND (ITEM5=POOR) AND (ITEM25>POOR) AND (ITEM12>POOR)	FAIR	66.7
20	IF (ITEM15=POOR) AND (ITEM5=POOR) AND (ITEM25>POOR) AND (ITEM12=POOR) AND (ITEM10>POOR)	FAIR	55.8
10	IF (ITEM15>POOR) AND (ITEM12<=FAIR) AND (ITEM4>POOR)	FAIR	84.1
16	IF (ITEM15>POOR) AND (ITEM12<=FAIR) AND (ITEM4=POOR) AND (ITEM12>POOR)	FAIR	81,9
22	IF (ITEM15>POOR) AND (ITEM12<=FAIR) AND (ITEM4=POOR) AND (ITEM12=POOR) AND (ITEM22>POOR)	FAIR	70.9
28	IF (ITEM15>POOR) AND (ITEM12<=FAIR) AND (ITEM4=POOR) AND (ITEM12=POOR) AND (ITEM22=POOR) AND (ITEM25>POOR) AND (ITEM14>POOR)	FAIR	63.5
11	IF (ITEM15>POOR) AND (ITEM12=GOOD) AND (ITEM25<=FAIR)	FAIR	74.9
17	IF (ITEM15>POOR) AND (ITEM12=GOOD) AND (ITEM25=GOOD) AND (ITEM11<=FAIR)	FAIR	60.0
23	IF (ITEM15>POOR) AND (ITEM12=GOOD) AND (ITEM25=GOOD) AND (ITEM11=GOOD) AND (ITEM2<=FAIR)	FAIR	55.4
24	IF (ITEM15>POOR) AND (ITEM12=GOOD) AND (ITEM25=GOOD) AND (ITEM11=GOOD) AND (ITEM2=GOOD)	GOOD	69.8

increases. Other variables also conduct to a POOR overall quality, as 'Fare/Service Ratio' when it is stated with POOR satisfaction. Finally, it can be seen that the precision values of these rules are quite high, the minimum value of them being 55.4% and the maximum 84.1%.

Another important result obtained in this model is the importance of the variables. This is achieved by using the importance index (Kashani, Mohaymany 2011), of which a standardized form has been used in this paper.

Table 4 shows the standardized importance rates stated by the users as well as the standardized importance rates deduced by the model.

By analysing these results, large differences can be observed. The most important items identified by the model concern Comfort, Personnel, Information and Service factors. This matches the results of another recent study analysing the same rail services (Eboli, Mazzulla 2012b), in which also these factors have been detected as important. The items with the highest importance deduced by the model are: ITEM11 'Windows and Doors Working', ITEM25 'Courtesy and Competence on Board', ITEM21 'Clear and fast info in the stations', ITEM14 'Train Punctuality', ITEM27 'Courtesy and Competence in Station' and ITEM15 'Regularity of Runs', all of them with standardized importance values exceeding 64.6%. Little importance has been deduced for the items related to Safety, Cleanliness or Other.

On the other hand, the most important factors stated by the users are the items concerning safety (*Item* 1, *Item* 2 and *Item* 3), which are not really relevant for the model (standardized importance values lower than 41%). However, when users are asked to rate the importance of each attribute, they consider all the attributes as highly important, as reported in the previous section. This is one of the serious drawbacks encountered when studying the importance of variables based on the stated opinions of passengers (Weinstein 2000).

Conclusions

Many studies analysed rail service quality using different approaches; however, this is the first time that CART methodology was applied for this purpose. In this paper, it was demonstrated that this new approach can successfully manage this problem, predicting the overall quality of the service with a high precision rate (in this

Table 4. Importance rates

Stated Importance			Derived Importance		
ITEM3	Personal Security at Station	100%	ITEM11	Windows and Doors Working	100%
ITEM2	Personal Security on Board	99.9%	ITEM25	Courtesy and Competence on Board	88.5%
ITEM1	Travel safety	99.7%	ITEM21	Information at Stations	81.0%
ITEM14	Punctuality of Runs	97.8%	ITEM14	Punctuality of Runs	73.5%
ITEM20	Facilities for Disabled	96.8%	ITEM27	Courtesy and Competence in Station	69.3%
ITEM15	Regularity of Runs	96.4%	ITEM15	Regularity of Runs	64.6%
ITEM12	Fare/Service Ratio	95.5%	ITEM12	Fare/Service Ratio	55.4%
ITEM5	Cleanliness of Seats	95.3%	ITEM13	Frequency of Runs	54.8%
ITEM13	Frequency of Runs	95.1%	ITEM1	Travel safety	41.0%
ITEM6	Cleanliness of Toilet Facilities	94.8%	ITEM17	Localization of Stations	40.9%
ITEM4	Cleanliness of Vehicles	94.7%	ITEM19	Bicycle Transport on Board	39.1%
ITEM21	Information at Stations	92.0%	ITEM10	Air-conditioning on Board	37.6%
ITEM22	Information on Board	91.3%	ITEM16	Price Integration with PT	37.6%
ITEM10	Air-conditioning on Board	90.9%	ITEM22	Information on Board	35.4%
ITEM16	Price Integration with PT	90.1%	ITEM4	Cleanliness of Vehicles	35.2%
ITEM27	Courtesy and Competence in Station	89.9%	ITEM5	Cleanliness of Seats	33.4%
ITEM25	Courtesy and Competence on Board	89.9%	ITEM2	Personal Security on Board	28.3%
ITEM17	Localization of Stations	89.9%	ITEM3	Personal Security at Station	28.1%
ITEM11	Windows and Doors Working	89.9%	ITEM7	Cleanliness of Stations	25.8%
ITEM7	Cleanliness of Stations	89.7%	ITEM8	Maintenance of Stations	25.4%
ITEM9	Crowding on Board	89.7%	ITEM23	Complaints	19.3%
ITEM23	Complaints	89.6%	ITEM24	Info Connections with PT	19.3%
ITEM24	Info Connections with PT	88.5%	ITEM26	Ticket Inspection	15.8%
ITEM8	Maintenance of Stations	88.0%	ITEM6	Cleanliness of Toilet Facilities	10.6%
ITEM26	Ticket Inspection	87.6%	ITEM20	Facilities for Disabled	9.1%
ITEM18	Parking	85.9%	ITEM18	Parking	7.1%
ITEM19	Bicycle Transport on Board	79.1%	ITEM9	Crowding on Board	6.8%

case 77.43%, which is higher than in other CART studies with similar objectives).

Moreover, CART methodology provides an alternative to parametric models. It has multiple advantages over the ordinary statistical modelling techniques. CART analysis allows using many explanatory variables. In this research 27 different variables describing railway service characteristics were used, and the most important variables were easily identified: 'Windows and doors working on board', 'Courtesy and Competence on Board', 'Information at Stations', 'Train Punctuality', 'Courtesy and Competence in Station' and 'Regularity of Runs'. The most important variables deduced by the model differ from the ones stated by the users in the survey.

Another important advantage of the CART model is that the outcomes of the analysis are easily understandable because they are represented in visual branching images. It also extracts decision rules providing useful information to public transport managers and operators in order to prioritize the measures that are going to develop in the service. Some conclusion about these rules can be extracted:

- a POOR satisfaction with 'Regularity of Runs' is the main cause for a POOR overall SQ;
- the probability of having a POOR perception of the overall SQ increases when the 'Cleanliness of Seats' has a POOR satisfaction (*Node* 3);
- also this probability increases when the satisfaction with the 'Fare/Service Ratio' is POOR (*Node* 13);
- on the other hand, when 'Regularity of Runs' is not POOR, a GOOD 'Fare/Service Ratio' and a GOOD 'Courtesy and Competence on Board' increase the probability of having a GOOD perception of the overall SQ.

Another important advantage of the CART model is that it does not need to establish a functional relationship among variables as ordinary statistical modelling techniques, such as regression models. Also, it can effectively handle multi-collinearity problems. The existence of multi-collinearity is very frequent in these satisfaction surveys. In regression analysis if the model is misspecified or exists multi-collinearity, the estimated relationship between dependent variable and independent variables as well as model predictions will be erroneous.

Not all are benefits, however. The classification tree models are generally 'unstable' because the building of the trees is based on their seed number, which is random, and therefore different trees could be obtained and the results might vary. This is the reason why tree models are often used only to identify important variables and other modelling techniques are used to develop final models.

References

Abdel-Aty, M.; Keller, J.; Brady, P. A. 2005. Analysis of types of crashes at signalized intersections by using complete crash data and tree-based regression, *Transportation Research Record* 1908: 37–45. http://dx.doi.org/10.3141/1908-05

Abreha, D. A. 2007. *Analysing Public Transport Performance Using Efficiency Measures and Spatial Analysis; the Case Study of Addis Ababa, Ethiopia:* Ms Thesis. International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands. 95 p. Available from Internet: http://www.itc.nl/library/papers_2007/msc/upla/abreha.pdf

Badami, M. G.; Haider, M. 2007. An analysis of public bus transit performance in Indian cities, *Transportation Research Part A: Policy and Practice* 41(10): 961–981. http://dx.doi.org/10.1016/j.tra.2007.06.002

Breiman, L.; Friedman, J. H.; Olshen, R. A.; Stone, C. J. 1984. *Classification and Regression Trees*. Chapman and Hall/CRC. 368 p.

Cavana, R. Y.; Corbett, L. M.; Lo, Y. L. 2007. Developing zones of tolerance for managing passenger rail service quality, *International Journal of Quality & Reliability Management* 24(1): 7–31. http://dx.doi.org/10.1108/02656710710720303

Chang, L.-Y.; Chen, W.-C. 2005. Data mining of tree-based models to analyze freeway accident frequency, *Journal of Safety Research* 36(4): 365–375. http://dx.doi.org/10.1016/j.jsr.2005.06.013

Chang, L.-Y.; Wang, H.-W. 2006. Analysis of traffic injury severity: an application of non-parametric classification tree techniques, *Accident Analysis & Prevention* 38(5): 1019–1027. http://dx.doi.org/10.1016/j.aap.2006.04.009

Chen, W.-H.; Jovanis, P. P. 2000. Method for identifying factors contributing to driver-injury severity in traffic crashes, *Transportation Research Record* 1717: 1–9. http://dx.doi.org/10.3141/1717-01

Cirillo, C.; Eboli, L.; Mazzulla, G. 2011. On the asymmetric user perception of transit service quality, *International Journal of Sustainable Transportation* 5(4): 216–232. http://dx.doi.org/10.1080/15568318.2010.494231

Council, F. M.; Stewart, J. R. 1996. Severity indexes for roadside objects, *Transportation Research Record* 1528: 87–96. http://dx.doi.org/10.3141/1528-09

De Oña, J.; De Oña, R.; Calvo, F. J. 2012. A classification tree approach to identify key factors of transit service quality, *Expert Systems with Applications* 39(12): 11164–11171. http://dx.doi.org/10.1016/j.eswa.2012.03.037

De Oña, J.; De Oña, R.; Eboli, L.; Mazzulla, G. 2013. Perceived service quality in bus transit service: a structural equation approach, *Transport Policy* 29: 219–226. http://dx.doi.org/10.1016/j.tranpol.2013.07.001

Dell'Olio, L.; Ibeas, A.; Cecín, P. 2011. The quality of service desired by public transport users, *Transport Policy* 18(1): 217–227. http://dx.doi.org/10.1016/j.tranpol.2010.08.005

Eboli, L.; Mazzulla, G. 2012a. Performance indicators for an objective measure of public transport service quality, *European Transport – Trasporti Europei* 51: 1–21.

Eboli, L.; Mazzulla, G. 2012b. Structural equation modelling for analysing passengers' perceptions about railway services, *Procedia – Social and Behavioral Sciences* 54: 96–106. http://dx.doi.org/10.1016/j.sbspro.2012.09.729

Eboli, L.; Mazzulla, G. 2011. A methodology for evaluating transit service quality based on subjective and objective measures from the passenger's point of view, *Transport Policy* 18(1): 172–181. http://dx.doi.org/10.1016/j.tranpol.2010.07.007

Eboli, L.; Mazzulla, G. 2010. How to capture the passengers' point of view on a transit service through rating and choice options, *Transport Reviews* 30(4): 435–450. http://dx.doi.org/10.1080/01441640903068441

Eboli, L.; Mazzulla, G. 2008a. A stated preference experiment for measuring service quality in public transport, *Transportation Planning and Technology* 31(5): 509–523. http://dx.doi.org/10.1080/03081060802364471

Eboli, L.; Mazzulla, G. 2008b. Willingness-to-pay of public transport users for improvement in service quality, *European Transport – Trasporti Europei* 38: 107–118.

Gini, C. 1912. *Variabilità e mutabilità*. Tipogr. di P. Cuppini. 158 p. (in Italian).

Habib, K. M. N; Kattan, L.; Islam, T. 2011. Model of personal attitudes towards transit service quality, *Journal of Advanced Transportation* 45(4): 271–285. http://dx.doi.org/10.1002/atr.106

Hallmark, S. L.; Guensler, R.; Fomunung, I. 2002. Characterizing on-road variables that affect passenger vehicle modal operation, *Transportation Research Part D: Transport and Environment* 7(2): 81–98. http://dx.doi.org/10.1016/S1361-9209(01)00012-8

Hensher, D. A.; Prioni, P. 2002. A service quality index for area-wide contract performance assessment, *Journal of Transport Economics and Policy* 36(1): 93–113.

Hensher, D. A.; Stopher, P.; Bullock, P. 2003. Service quality – developing a service quality index in the provision of commercial bus contracts, *Transportation Research Part A: Policy and Practice* 37(6): 499–517. http://dx.doi.org/10.1016/S0965-8564(02)00075-7

Huang, C.-F.; Hsueh, S.-L. 2010. Customer behavior and decision making in the refurbishment industry – a data mining approach, *Journal of Civil Engineering and Management* 16(1): 75–84. http://dx.doi.org/10.3846/jcem.2010.07

Jen, W.; Tu, R.; Lu, T. 2011. Managing passenger behavioral intention: an integrated framework for service quality, satisfaction, perceived value, and switching barriers, *Transportation* 38(2): 321–342. http://dx.doi.org/10.1007/s11116-010-9306-9

Joewono, T. B.; Kubota, H. 2007. User perception of private paratransit operation in Indonesia, *Journal of Public Transportation* 10(4): 99–118.

Kashani, A. T.; Mohaymany, A. S. 2011. Analysis of the traffic injury severity on two-lane, two-way rural roads based on classification tree models, *Safety Science* 49(10): 1314–1320. http://dx.doi.org/10.1016/j.ssci.2011.04.019

Kuhnert, P. M.; Do, K.-A.; McClure, R. 2000. Combining nonparametric models with logistic regression: an application to motor vehicle injury data, *Computational Statistics & Data Analysis* 34(3): 371–386. http://dx.doi.org/10.1016/S0167-9473(99)00099-7

Lao, Y.; Liu, L. 2009. Performance evaluation of bus lines with data envelopment analysis and geographic information systems, *Computers, Environment and Urban Systems* 33(4): 247–255. http://dx.doi.org/10.1016/j.compenvurbsys.2009.01.005

Magazzù, D.; Comelli, M.; Marinoni, A. 2006. Are car drivers holding a motorcycle licence less responsible for motorcycle – car crash occurrence?: a non-parametric approach, *Accident Analysis & Prevention* 38(2): 365–370. http://dx.doi.org/10.1016/j.aap.2005.10.007

Nathanail, E. 2008. Measuring the quality of service for passengers on the Hellenic railways, *Transportation Research Part A: Policy and Practice* 42(1): 48–66. http://dx.doi.org/10.1016/j.tra.2007.06.006

Nocera, S. 2011. The key role of quality assessment in public transport policy, *TEC – Traffic, Engineering and Control* 52(9): 394–398.

Nocera, S.; Cavallaro, F. 2011. Policy effectiveness for containing CO_2 emissions in transportation, *Procedia – Social and Behavioral Sciences* 20: 703–713. http://dx.doi.org/10.1016/j.sbspro.2011.08.078

Nocera, S.; Cavallaro, F. 2012. Economic evaluation of future carbon dioxide impacts from Italian highways, *Procedia – Social and Behavioral Sciences* 54: 1360–1369. http://dx.doi.org/10.1016/j.sbspro.2012.09.850

Pakdil, F.; Aydın, Ö. 2007. Expectations and perceptions in airline services: an analysis using weighted SERVQUAL scores, *Journal of Air Transport Management* 13(4): 229–237. http://dx.doi.org/10.1016/j.jairtraman.2007.04.001

Pande, A.; Abdel-Aty, M.; Das, A. 2010. A classification tree based modeling approach for segment related crashes on multilane highways, *Journal of Safety Research* 41(5): 391–397. http://dx.doi.org/10.1016/j.jsr.2010.06.004

Qin, X.; Han, J. 2008. Variable selection issues in tree-based regression models, *Transportation Research Record* 2061: 30–38. http://dx.doi.org/10.3141/2061-04

SAS Institute Inc. 2004. *SAS/STAT® 9.1 User's Guide*. SAS Publishing. Cary, NC, U.S. 5136 p. Available from Internet: http://support.sas.com/documentation/onlinedoc/91pdf/sasdoc_91/stat_ug_7313.pdf

Sheth, C.; Triantis, K.; Teodorović, D. 2007. Performance evaluation of bus routes: a provider and passenger perspective, *Transportation Research Part E: Logistics and Transportation Review* 43(4): 453–478. http://dx.doi.org/10.1016/j.tre.2005.09.010

Sohn, S. Y.; Shin, H. 2001. Pattern recognition for road traffic accident severity in Korea, *Ergonomics* 44(1): 107–117. http://dx.doi.org/10.1080/00140130120928

TRB. 2003. *A Summary of TCRP Report 88: A Guidebook for Developing a Transit Performance-Measurement System*. Transportation Research Board (TRB), National Research Council, U.S. 24 p. Available from Internet: http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_report_88/SummaryDoc.pdf

Wang, I.-M.; Shieh, C.-J. 2006. The relationship between service quality and customer satisfaction: the example of CJCU library, *Journal of Information and Optimization Sciences* 27(1): 193–209. http://dx.doi.org/10.1080/02522667.2006.10699686

Washington, S.; Wolf, J. 1997. Hierarchical tree-based versus ordinary least squares linear regression models: theory and example applied to trip generation, *Transportation Research Record* 1581: 82–88. http://dx.doi.org/10.3141/1581-11

Washington, S.; Wolf, J.; Guensler, R. 1997. Binary recursive partitioning method for modeling hot-stabilized emissions from motor vehicles, *Transportation Research Record* 1587: 96–105. http://dx.doi.org/10.3141/1587-11

Weinstein, A. 2000. Customer satisfaction among transit riders: how customers rank the relative importance of various service attributes, *Transportation Research Record* 1735: 123–132. http://dx.doi.org/10.3141/1735-15

Wong, J.-Y.; Chung, P.-H. 2007. Managing valuable Taiwanese airline passengers using knowledge discovery in database techniques, *Journal of Air Transport Management* 13(6): 362–370. http://dx.doi.org/10.1016/j.jairtraman.2007.07.001

Yan, X.; Radwan, E. 2006. Analyses of rear-end crashes based on classification tree models, *Traffic Injury Prevention* 7(3): 276–282. http://dx.doi.org/10.1080/15389580600660062

Yan, X.; Richards, S.; Su, X. 2010. Using hierarchical tree-based regression model to predict train–vehicle crashes at passive highway-rail grade crossings, *Accident Analysis & Prevention* 42(1): 64–74. http://dx.doi.org/10.1016/j.aap.2009.07.003