Transportation Research Part A xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Spatial variation of the perceived transit service quality at rail stations

Laura Eboli, Carmen Forciniti, Gabriella Mazzulla*

University of Calabria, Department of Civil Engineering, P. Bucci, cubo 46/B, 87036 Rende, CS, Italy

ARTICLE INFO

Keywords:
Transit service quality
Rail stations
Passengers' perception
Spatial variation

ABSTRACT

Passenger's perception about transit service quality strongly depends on the context to which the passenger refers. Consequently, each attribute affecting service quality and rated in terms of satisfaction or importance could show spatial pattern across the region and/or temporal variations. Although transit service quality was largely investigated, few studies took into account the spatial variability of passengers' responses, given that a large amount of previous studies considered customer satisfaction with transit services in a cross sectional manner only. In order to cover this lack, we propose a study focused on transit service quality evaluated at railway stations, aimed to suggest a methodology for treating spatial variation of service quality attributes. The research is supported by Customer Satisfaction Survey data collected by a railway operator providing regional services in the North of Italy, and particularly rail services of regional and suburban lines connecting different towns of the hinterland of the city of Milan. The emerging results show spatial patterns across the region of the transit service quality attributes which should be taken into account for a deeper investigation of passengers' perceptions about rail services.

1. Introduction

The assessment of transit service quality is a matter of academic interest since over thirty years already, and several research studies highly covered the scientific literature since the beginning of the new millennium.

The growing interest for this topic also from transit enterprises and authorities is due to the awareness that the knowledge of the transit service performances can be used for monitoring the services and for identifying potential weak points, with the aim to increase passenger's satisfaction and finally attract new transit users and retaining the usual ones.

In fact, it is well known that in recent years users' travel behaviour has become more complex, and people prefer private car for moving, because it is perceived as the most comfortable, flexible, and fast mode compared with the other ones. The urban sprawl and low density areas rising also contributed to increase the attractiveness of car as the most convenient transportation mode from all the points-of-view, with undesirable consequences in terms of environmental pollution, energy consumption and quality of life (Nocera and Cavallaro, 2016; Ayyildiz et al., 2017). In order to reduce these problems, travel demand has to be moved from private car to more sustainable mode as public transit systems. Consequently, transport companies have to operate interventions aimed to reduce travel time and increase service quality over all the aspects, since improvements in service quality can have a significant positive impact on the willingness of potential users to choose public over private transport (de Oña et al., 2014; Eboli et al., 2016).

Transit service quality can be measured by different approaches. The most common approach is based on transit users' opinions

* Corresponding author.
E-mail addresses: laura.eboli@unical.it (L. Eboli), carmen.forciniti@unical.it (C. Forciniti), gabriella.mazzulla@unical.it (G. Mazzulla).

https://doi.org/10.1016/j.tra.2018.01.032

about the used services, collected by means of surveys. When service quality is measured from the customer's perspective, transit quality depends on the passengers' perceptions about each attribute characterising the service (de Oña et al., 2013). In recent years, the most popular methods for capturing transit users' perceptions were: (1) traditional Customer Satisfaction Surveys (CSS), where users express their opinions by rating the various service characteristics (Weinstein, 2000; Cavana et al., 2007; Joewono and Kubota, 2007; Pakdil and Aydın, 2007; Dell'Olio et al., 2011; Eboli and Mazzulla, 2010, 2012b; Jen et al., 2011; Habib et al., 2011); and (2) Stated Preference surveys, where the importance given by the users to the service attributes indirectly derives from 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 and Mazzulla, 2008a, 2008b, 2010; Hensher and Prioni, 2002; Hensher et al., 2003; Dell'Olio et al., 2011).

An alternative approach entails using different variables of the transit system demand and operation to calculate the 'efficiency' indicators (e.g. Badami and Haider, 2007; Lao and Liu, 2009; Eboli and Mazzulla, 2012a). There are also some studies integrating the two mentioned approaches (e.g. Sheth et al., 2007; Abreha, 2007; Nathanail, 2008; Eboli and Mazzulla, 2011).

Overall quality of transit services contains a large number of criteria representing the customer point-of-view about the service provided. As an example, the European Standard EN13816 (CEN, 2002) grouped the criteria into eight categories as: availability, accessibility, information, time, customer care, comfort, security, environmental impact; for each category different sub-levels are established. With the exception of the first and second categories, which describe transit services in a more general terms (e.g. extent of the service provided in terms of geography, time and frequency; interface with other transport mode), the other categories contain specific aspects which can refer to on-board service characteristics, at stations or stops service characteristics, or both of them. As an example, by considering time criterion, we can have on-board travel time, service punctuality, service regularity, but also access/egress time. Information criterion can be seen as travel information on-board and also information at stations or stops. Analogously, customer care criterion distinguishes between staff appearance/behaviour by referring separately to on-board staff and staff at stations/stops and, in the same way, comfort criterion can be considered in terms of ride comfort or cleanliness on-board or at stations/stops. Consequently, certain aspects should be referred to a specific line or route, whereas others should be referred to a specific station or stop.

Although transit service quality was largely investigated, only few studies considered the variability of passengers' opinions in terms of space and time, given that the major part of previous studies analysed passengers' satisfaction with transit services in a cross sectional manner only. However, it is common knowledge that passenger's perceptions about transit service quality attributes strongly depend on the context to which the passenger refers. For this reason, each attribute judged by users in terms of satisfaction or importance could show spatial and/or temporal variability.

Some authors pointed out transit service quality temporal variability. As an example, Cats et al. (2015) investigated the changes in the attributes' importance level over a more than ten years' time span for an average Swedish transit user. They confirmed that the perceived importance varies considerably over time for some service quality attributes. de Oña et al. (2016) proposed the use of the index numbers usually applied in the economic and industrial field for analysing the variation of service quality over time, by providing a powerful measurement for making comparisons of passengers' satisfaction with public transport in Spain on different temporal periods.

Other authors analysed the spatial variation of passengers' satisfaction with transit services by distinguishing among the different routes or lines (see for example Nathanail, 2008; Diab and El-Geneidy, 2012). On the other hand, some researchers related passengers' perceptions to a specific stop or station (see for example Iseki and Taylor, 2010; Hernandez et al., 2016). Nkurunziza et al. (2012) integrated random utility models into geo-spatial models; more specifically, they derived random utility functions from stated choice modelling and, by using a geographic information system (GIS), spatially analysed the potential user preferences and identified the preference variation among zones. Grisè and El-Geneidy (2017) presented a new method to spatially evaluate CSS data examining satisfaction with bus service across various neighbourhoods of London. They adopted multi-level regression modelling to estimate the relationship between overall satisfaction and social deprivation of the area where bus routes were operating.

Few researches focused on the analysis of both spatial and temporal variations. As an example, Diab and El-Geneidy (2014) tried to better understand transit passengers' perception towards the implementation of various improvement strategies in bus service over time. The study analysed Canadian bus user perceptions over a period of three years, by using stop-level data collected from automatic vehicle location (AVL) and automatic passenger count (APC) systems. Descriptive statistics and regression models were adopted to help in better understanding the differences between perceptions and reality. Abenoza et al. (2017) characterized current and potential users of public transport in Sweden and explored whether service attributes' importance changes over time for different segments of travellers. In addition, they investigated the variation of overall satisfaction among different geographical regions. To this end, year-based ordered logit models were estimated for each of the traveller segments.

The previous studies were based on approaches that do not denote a deep manner to analyse spatial variation intrinsically due to the characteristics and features of the region where the transit system operates. In fact, all these approaches were based on aggregate models which partially considered the spatial and/or temporal variations of the overall satisfaction with the services, by ignoring the geographic variations of each specific service quality attribute across the space.

We retain that by using spatial analysis techniques we can found the potential presence of spatial patterns of the transit service quality attributes. For investigating on spatial variations of service quality in a region, responses of transit users expressed by CSS have to be related to a specific location in the space, as user's neighbourhood, bus stop or railway station, which are origin and destination of the trip. In this way, passengers' perceptions can be related to the stop or station.

The present paper focuses on transit service quality evaluated at railway stations, and aims to suggest a methodology for analysing spatial variation of service quality attributes. The study is supported by CSS data concerning a railway operator offering regional services in the North of Italy, and particularly regional and suburban rail services operating in the hinterland of the city of Milan.

After selecting the service attributes which refer specifically to the railway stations, data are analysed using techniques of spatial statistics and modelled by means of a spatial regression model, named geographically regression model (GWR), in order to investigate on the spatial variation of transit service quality attributes across the study area.

This paper begins with a section describing the proposed methodology. Section 3 regards the study case and the experimental context. In Section 4, we declare the results obtained from the application of the proposed methodology with particular regard to the spatial association and GWR techniques. Finally, Section 5 contains discussion about results and final remarks.

2. Methodology

In order to investigate on spatial variation of transit service quality attributes, data have to be analysed for evaluating the potential presence of spatial patterns.

Generally, the studies that investigate on the quality of transit service ignore the geographic variations of attributes of service quality across the space. This implies the violation of the basic assumptions of independence and homogeneity implicit in conventional statistical analysis and could lead to not satisfactory results. Assuming the hypothesis of stationary distribution of the variables leads to information loss, biased and/or inefficient parameters and the possibility of seriously flawed conclusions and policy prescriptions (Griffith and Layne, 1999).

The proposed methodology develops in a few steps described in the following.

Step 0. Selecting attributes of service quality

A group of service quality attributes eligible for spatial analysis is selected among the aspects referring to rail stations. The selection is made by excluding from the analysis the service aspects that are not directly linked to the station.

Step 1. Analysing preliminarily passengers' perceptions

The perceptions expressed by the passengers through the questionnaire are analysed by basic descriptive statistics such as average values, frequencies distribution, skewness and kurtosis. These data are mapped by means of GIS for having a first qualitative visualization of the spatial variation of the attributes of service quality.

Step 2. Defining spatial units

A spatial allocation is associated to each interview by using the information about place of residence, access and egress stations collected through the questionnaire. Each interviewed passenger is linked with a municipality on the basis of the place of residence or station for accessing the rail services. Then, each municipality is assigned to the nearest station; in some cases, more municipalities are assigned to the same station because there are various municipalities without a rail station inside their boundary. Definitively, the spatial unit represents the municipality or a group of municipalities associated to one rail station. We adopt this unit of analysis because we retain that the agglomeration of municipalities better represents spatially the locations where respondents live.

Step 3. Identifying global spatial autocorrelation among the rating values

For identifying the possible presence of spatial variability in the data, particular statistical tests, called statistics of spatial autocorrelation, can be performed. We use autocorrelation for verifying if in the spatial distribution of the data there is aggregation of similar values. In this case, the values of the variable present a spatial variation that is systematic and not random. Spatial autocorrelation is the tendency of variables to display some degree of systematic spatial variation. This fact often means that data from locations near to each other are usually more similar than data from locations far away from each other.

Spatial autocorrelation can be detected performing the exploratory spatial data analysis (ESDA), which is a collection of techniques to visualise spatial distributions, identify atypical locations or spatial outliers, discover patterns of spatial association, clusters or hot spots, and suggest spatial regimes or other forms of spatial heterogeneity. Among the most familiar tests for global spatial autocorrelation, we choose to use Getis-Ord General G (Getis and Ord, 1992) and Moran's *I* (Anselin et al., 2007).

Step 4. Identifying local spatial autocorrelation among the rating values

After detecting spatial autocorrelation by means of global statistics, places with unusual concentrations of high or low values can be detected using local statistics of spatial association. We decided to use the two statistics which have been mostly used in the literature: $G_i(d)$ statistics (Ord and Getis, 1995, 2001) and Local Indicators of Spatial Association (LISA) as Local Moran's *I* (Anselin, 1995).

Step 5. Evaluating average weights of service quality attributes

If the variables present spatial autocorrelation, data could be treated through the techniques of spatial analysis. For this purpose, a regression analysis represents a valid tool. Specifically, a linear regression model, as Ordinary Least Square (OLS), can be useful for checking if the residuals of the regression are autocorrelated. In addition, OLS permits to estimate an average value of the coefficients for the entire study area for having a measure of the weight of each coefficient.

However, developing OLS using data having spatial patterns could show some lacks, because ignoring the geographic variations of variables represents a relevant disadvantage. In fact, an assumption of OLS analysis is the existence of stationary relationships among the variables ignoring local variation caused by spatial heterogeneity (Paéz, 2006). In this case, GWR must be performed.

Step 6. Evaluating spatial variation of service quality attributes by spatial regression models

Geographically Weighted Regression (GWR) are performed to model spatially varying relationships. GWR is a local form of linear regression and constructs a separate equation for every spatial unit in which the study area is divided.

As reported by Cardozo et al. (2012), the use of GWR is supported by many advantages. First of all, GWR results provide greater detail and accuracy moving from a global perspective to a local analysis (Lloyd, 2010). Another important advantage of spatial modelling is measuring spatial instability from the magnitude of the coefficient across the area (Clark, 2007). This is not possible with traditional regression models because they assume that the coefficients have not significant differences in space (Cardozo et al., 2012). It is possible to analyse how relationships vary over space and to investigate the spatial pattern of the local estimates (Fotheringham et al., 2002). In addition, the analysis of the spatial distribution of the local coefficients allows to recognize where the independent variables have greater or lesser explanatory power. Hadayeghi et al. (2010) explain that when using GWR, in most of the cases, estimation errors are smaller and the problem of spatial autocorrelation is eliminated or reduced. GWR has also practical application, because it provides specific results for each location that can be used as evidence to support local policies and decision-making. For this reason, these techniques of local regression are frequently referred to as "place-based" (de Smith et al., 2009).

3. Experimental context

3.1. Data

The data supporting this research were collected through a survey conducted by a railway operator providing regional services in the North of Italy, and addressed to its passengers. The analysed service offers different types of connections, such as regional lines and suburban lines connecting towns of the hinterland of Milan.

About 570,000 passengers per day travel by the analysed lines; as an example, a regional line is used by about 25,000/30,000 passengers per day. The maximum length of a line is about 150 km (e.g. the lines connecting the city of Milan and the hinterland 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 conducted through a questionnaire addressed to a sample of about 40,000 users in consecutive survey campaigns between May 2013 and May 2014. The interviews were conducted on board during the whole week (even in holiday days) and at any hour of the day. The questionnaire was designed into two sections. The first one regards general information about the trip, user's socio-economic characteristics and travel habits. By this section, information about the station of access to train and egress from train were collected; this information allowed to associate an access station to each interview.

The second section focuses on passengers' perceptions of the used services. Users were asked to express importance and satisfaction rates about service quality factors revealed on board and at the station where the trip starts. Rates were given on an ordinal scale from 1 to 10. An amount of 33 attributes of service quality were investigated, classified into seven categories, as safety, cleanliness, comfort, service, additional services, information, and personnel.

In this study, seven attributes were chosen because they are the sole attributes that concern aspects for which service quality can depend on station, that is it can vary over the space. Table 1 shows the list of the chosen attributes with the average rates of importance and satisfaction.

From Table 1, we can highlight that "Personal security at station" and "Facilities for disabled" are the attributes with the highest importance rates, whereas "Maintenance of station" is the attribute evaluated with the lowest importance rate, although it received a relevant rate, higher than 8. On the other hand, users expressed satisfaction rates that are much lower than the importance rates. More specifically, users result rather unsatisfied with attributes of service quality such as "Facilities for disabled" and "Info

Table 1
Importance and satisfaction rates of the analysed service quality attributes.

	Service quality attributes	Importance rate	Satisfaction rate
Safety	Personal security at station	9.17	6.01
Cleanliness	Cleanliness of station	8.31	5.14
	Maintenance of station	8.13	5.14
Service	Integration with other transit systems	8.43	5.82
Additional services	Facilities for disabled	9.01	4.78
Information	Information at station	8.59	5.32
	Info connections with other transit systems	8.27	4.90

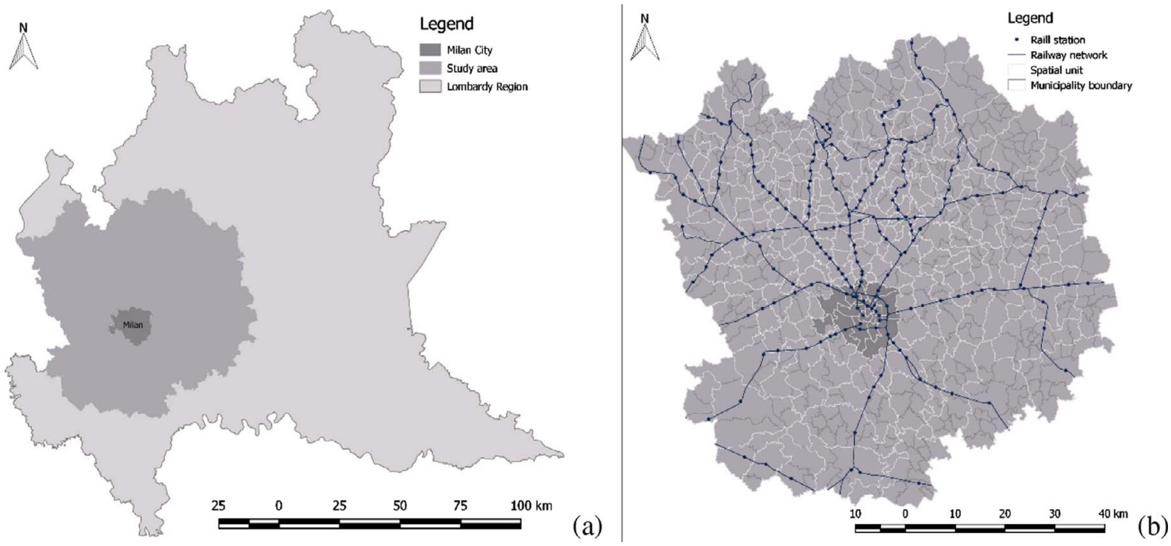


Fig. 1. Study area compared with the Lombardy region (a); railway network and stations in the study area (b).

connection with other transit systems". In addition, "Facilities for disabled" is the attributes that shows the greatest distance between the importance rate and the satisfaction one. Specifically, users give much importance to this attribute but, at the same time, they result not adequately satisfied. On the other hand, "Integration with other transit system" is the attribute with the lowest difference between importance and satisfaction rate.

3.2. Study area

Study area was defined by overlapping information from different sources, such as the number of interviews made at each railway station and the number of daily trips made by train in each municipality. Using a GIS, each municipality was assigned to the nearest station on the basis of the distance between the centroid of the town zone and the station. In many cases, more municipalities were associated to the same station. In this way, an area of influence was assigned to each railway station. In order to reduce the study area, the map of the number of interviews at stations and of the number of trips made by resident people in each municipality were overlapped to the areas of influence of the stations. As a result, in the study area only the stations with more than 5 interviews and more than 20 trips were included. The final study area resulted centred on the city of Milan. An amount of 674 municipalities was selected and assigned to 200 railway stations. As reported in Fig. 1, the study area has a circular shape and includes the radial railway routes that connect the city of Milan with the suburban areas and with other municipalities of the Lombardy region.

The total population is equal to 6,816,552 inhabitants; in average females are slightly more numerous (51.6%) than males (48.4%). Workers (resident people that have a job) correspond to 44.2% of total population.

Data collected during the survey were elaborated and assigned to each station, and then to each spatial unit in order to apply clustering techniques. The sample was limited to 23,015 interviews based on the selected study area. In this study, only the attributes of service quality evaluated in terms of satisfaction rates were considered.

Table 2 reports the frequencies and the descriptive statistics of the data referred to the stations, including skewness and kurtosis. A negative skew with a value lower than −0.50 represents a negative asymmetric distribution, in which the tail tends towards left; a value higher than 0.50 detects a positive asymmetric distribution, which is towards right; and a value ranging from -0.50 to 0.50 detects a distribution that can be considered as symmetric. All the analysed factors show values of skewness from −0.50 to 0.50; therefore the distributions can be considered as symmetric.

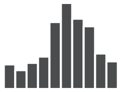
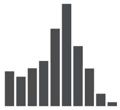
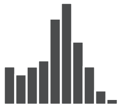
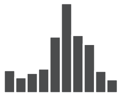
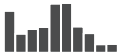
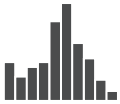
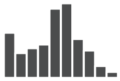
Kurtosis is a descriptor of the shape of a frequency distribution that provides a comparison of the shape of a given distribution to the normal one. Specifically, a positive kurtosis indicates a leptokurtic distribution (long tails and pronounced gibbosity), while a negative kurtosis indicates a platykurtic distribution (small tails and poor gibbosity); finally, if the value is around zero, the distribution is mesokurtic (normal).

All factors show a distribution with negative kurtosis, and the histograms show small tails and poor gibbosity. In some cases, as for the factor "Integration with other transit systems", it seems to have a normal distribution.

By analysing the distribution of the values for each attribute of service quality, it results that the values of mean and of median are lower than the values of mode. This means that the lowest rates of the attribute are more numerous, even if the value of the distribution having the highest frequency results higher. Although the value of the mode calculated for each factor is equal to 6, the amount of score 6 is sharper for some factors, as "Cleanliness of station" and "Maintenance of station", than for other, as "Facilities for disabled". As an example, we can see that the histogram related to "Cleanliness of station" shows a higher bar corresponding to the 6 scores, compared to the rest of scores; otherwise, the histogram of "Facilities for disabled" includes bars having similar height.

In many cases, we can note that the frequency distributions of the rates present a high value of the lower rates, especially in the

Table 2
Descriptive statistics of the analysed service quality attributes.

Factors	Frequencies					Mean	Mode	Median	Skewness	Kurtosis	Histogram
	Tot	1–3	4–5	6–7	8–10						
Personal security at station	22,126	3403	4876	7673	6174	6.01	6	6	−0.369	−0.401	
Cleanliness of station	22,613	5303	6214	8157	2939	5.14	6	5	−0.249	−0.545	
Maintenance of station	22,501	5220	6378	8060	2843	5.14	6	5	−0.269	−0.516	
Integration with other transit systems	18,333	2899	3982	7269	4183	5.82	6	6	−0.438	−0.200	
Facilities for disabled	13,499	4199	3713	3771	1816	4.78	6	5	0.020	−0.761	
Information at station	21,912	4777	5827	7678	3630	5.32	6	6	−0.257	−0.504	
Info connections with other transit systems	17,695	4881	5030	5565	2219	4.9	6	5	−0.143	−0.684	

case of “Facilities for disabled” (31%), “Info connections with other transit system” (28%), “Cleanliness of station” (23%), “Maintenance of station” (23%).

In order to facilitate the graphic representation of data across the study area, and to obtain a unique value for each attribute to allocate to the station, we calculated the mode of the rates given by users that begin the trips from the same station. We made this assumption because the mode value allows maintaining the discrete framework of the rating scale used during the survey. The mode corresponds to the value of the distribution having the highest frequency. Consequently, we choose the mode for representing the satisfaction of service quality because it is not influenced by the amount of lower or higher values.

The values of mode were grouped into four classes. The first class (1–3) and the fourth class (8–10) are wider than the others, and include extreme values, whereas the second (4–5) and the third class (6–7) include the intermediate values. As shown in Fig. 2, the spatial distribution of data for each attribute displays the presence of spatial aggregations of similar values. By means of techniques of spatial statistics we can declare if the spatial pattern is the result of random choice or of spatial association.

By observing Fig. 2, we can say that the satisfaction with the various service aspects is quite variable over the space, and also that the variation is different according the specific service aspect. As an example, users of the major part of the spatial units are satisfied with the aspects concerning personal security, cleanliness, maintenance, and integration with other transit systems, while other aspects such as facilities of disabled and info connections with other transit systems are less satisfying for the users of many spatial units. Notwithstanding some aspects seem quite critical for several spatial units, users of almost all the spatial units gave a very positive judgment about the overall service, as we can observe in the last illustration of Fig. 2. In fact, overall service quality shows values prevalently included in third and fourth classes, between 6 and 10.

More specifically, the generalized presence of low values of judgements for “Facilities for disabled” and “Info connections with other transit systems” is due to an objective absence of adequate facilities for people with disabilities in a large part of the rail stations, and to the lack of information regarding the connections with other transit systems in such a stations more distant from the city of Milan. By analysing these areas the obtained results is not surprising. In fact, we talk about the southwest area (Vigevano), the southeast area (Crema), the northeast area (Bergamo), and sometimes the area north of Milan around the town of Lecco. In the last case, no criticism about personal security and integration with other transit systems were emerged. In the towns of Vigevano, Crema and Bergamo rail stations are located in peripheral districts, and therefore it is quite predictable that the level of degradation is greater than in the tows where the stations are located in the most central areas. As a consequence, it is expected that service quality attributes like cleanliness and maintenance of the stations, and personal security can have lower values of satisfactions by the

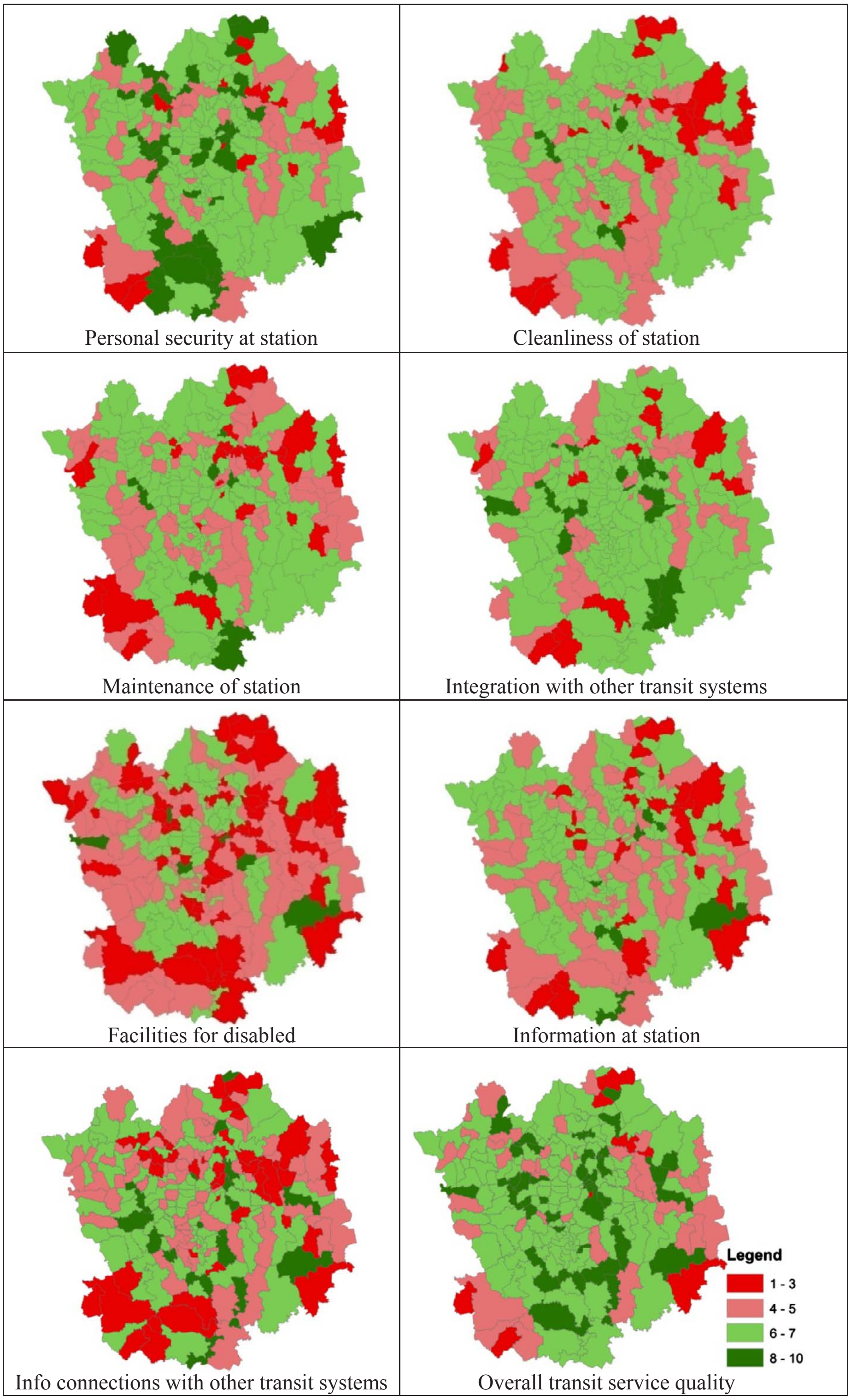


Fig. 2. Spatial variation of the attributes of transit service quality.

Table 3
Results of global spatial autocorrelation tests.

Variable	Getis-Ord General G				Moran's I Index			
	G Index	Z-score	p-Value	Pattern	I Index	Z-score	p-Value	Pattern
Personal security at station	0.000012	2.18	0.030	High cluster	0.0048	15.29	0.000	Clustered
Cleanliness of station	0.000012	2.94	0.003	High cluster	0.0100	31.78	0.000	Clustered
Maintenance of station	0.000012	2.13	0.034	High cluster	0.0085	27.16	0.000	Clustered
Integration with other transit systems	0.000012	2.15	0.032	High cluster	0.0025	8.20	0.000	Clustered
Facilities for disabled	0.000012	1.89	0.060	High cluster	0.0047	15.22	0.000	Clustered
Information at station	0.000012	2.03	0.043	High cluster	0.0029	9.43	0.000	Clustered
Info connections with other transit systems	0.000012	1.78	0.075	High cluster	0.0054	17.43	0.000	Clustered
Overall transit service quality	0.000012	-4.19	0.000	Lowcluster	0.0034	10.88	0.000	Clustered

passengers.

4. Results

4.1. Spatial association

Clustering techniques of spatial autocorrelation allow to determine whether or not inherent geographically based relationships exist. Global and local measures of spatial autocorrelation, as defined in Section 2, were applied on the attributes of transit service quality and implemented in a GIS environment.

At first, the tests for global spatial autocorrelation, Getis-Ord General G and Moran's Index I, were performed on data. Getis-Ord General G measures the degree of clustering for either high values or low values based on the attribute values of each couple of spatial units and the spatial weight between features. Getis-Ord General G was estimated by considering as spatial weight the inverse Euclidean distance among the spatial units. This statistical test begins by identifying the null hypothesis "there is no spatial clustering". The Z Score is a test of statistical significance for deciding whether or not to reject the null hypothesis, and is associated with a standard normal distribution. This distribution relates standard deviations with probabilities and allows significance and confidence to be attached to Z scores. The higher (or lower) the *Z* score, the stronger the intensity of the clustering. A *Z* score near zero indicates no apparent clustering within the study area. A positive *Z* score indicates clustering of high values; a negative *Z* score indicates clustering of low values.

As reported in Table 3, the results of Getis-Ord General G indicate the potential presence of clustering of high values for all the attributes; therefore, the distribution of values across the space has low likelihood to be the result of random choice.

Moran's *I* is essentially a cross product correlation measure that incorporates spatial weights and explains if data distribution is clustered or random. Moran's *I* was estimated also by considering as spatial weight the inverse Euclidean distance among the spatial units, and it serves as a test where the null hypothesis is the spatial independence (in this case its value would be zero). Positive values (between 0 and 1) indicate a direct correlation, and negative values (between -1 and 0) indicate an inverse correlation. To estimate the statistical significance of the index, a normal distribution has to be associated.

The results of Moran's *I* show another situation (Table 3). For all the attributes, the statistics assesses that the distributions of values are spatially clustered at the global scale.

The difference between the results provided by the two techniques of global spatial association can be explicated by considering the analytical formulation of each one. In fact, the Getis-Ord General G is calculated considering the values of the variable, whereas Moran's *I* is based on covariance. In addition, data included in this study are very disaggregated. As already specified, we considered an area served by 200 railway stations, and thus incorporating 200 spatial units.

As a consequence, in this case we can consider the Moran's Index *I* as the most appropriate index for detecting the potential presence of clustering for all the attributes. Definitively, based on the outcomes of global spatial statistics, we can assert that the distribution of the values of the service quality attributes present spatial autocorrelation. However, the local statistics of spatial association can be applied in order to localize the clusters of similar values. Specifically, the techniques named as Hot-spot analysis Getis-Ord and Anselin Local Moran's *I* were performed.

The Hot-spot analysis Getis-Ord statistics is a distance-based statistic and measures the proportion of a variable found within a given radius of a point, respective to the total sum of the variable in the study region. The statistics is able to individuate the potential presence of spatial agglomeration of relatively high values, or relatively low values clustered together.

The Anselin Local Moran's *I* allows for the decomposition of global indicators, such as Moran's *I*, into the contribution of each individual observation. Interpretation of the Local Moran's *I* is less intuitive than interpretation of the Hot-spot analysis Getis-Ord statistic. By means of a scatter plot, four patterns of local spatial association can be recognised:

1. High-high association: the value of a generic variable related to a spatial unit is above the mean, and the values of the same variable at neighboring zones are above the mean, the statistic is positive;
2. Low-low association: both values are below the mean, the statistic is positive;

3. High-low association: the value at a spatial unit is above the mean and the values at neighboring zones are below the mean, this gives a negative statistics;
4. Low-high association: the value at a spatial unit is below the mean and the weighted average is above the mean, Local Moran's *I* is negative.

The combination of Local Moran's *I* and the derived scatter plot tool provides information on different types of spatial association at the local level.

The tests were calculated by considering as spatial weight the inverse Euclidean distance among the spatial units. Both the tests give back as output the maps with the indication of the local spatial clusters (Table 4).

For individuating the potential presence of spatial agglomeration, the Z scores associated to the values of G for each spatial unit *i* (G_iZScore) were considered in Hot-spot analysis Getis-Ord. The statistics individuate the presence of spatial agglomeration of relatively high values clustered together when the values of G_iZScore are positive, and of relatively low values when the values of G_iZScore are negative.

The values of Anselin Local Moran's I individuate four patterns of local spatial association: high-high association (HH), low-low association (LL), high-low association (HL), low-high association (LH).

As shown in Table 4, the results of these two different spatial statistics display compliance respect the localization of clusters.

For each service quality attribute, with the exception of "Facilities for disabled", the results of local Getis-Ord and LISA show an aggregation of low values in the southwest area. This means that the area of Vigevano is characterized by low levels of service quality at station. "Facilities for disabled" was considered as a low quality factor in the northeast area (Lecco). At the same time, spatial agglomerations of relatively high values can be noticed in the central areas located at northwest area of Milan and bordering areas (corresponding to Rho and Saronno areas).

In addition, "Maintenance of station" shows agglomeration of high values near Pavia (southern area), and "Integration with other transit systems" presents clusters of high values in the areas corresponding to Monza (central area) and Lodi (southern area).

This analysis displays the presence of spatial patterns in attributes of transit service quality and localizes the aggregation of similar values across the study area.

4.2. Geographically weighted regression

Before applying GWR, good practice is performing OLS in order to prove the inappropriateness of no-spatial models for treating spatial data. The OLS and GWR models were elaborated using the same sample of 23,015 interviews and the same variables by means of ArcGIS (ESRI).

We run an OLS model for two main reasons. First, for checking if the residuals of the OLS regression are autocorrelated. In this case GWR must be performed. The second reason is to estimate an average value of the coefficients for the entire study area for having a measure of the weight of each coefficient.

The dependent variable is the passengers' satisfaction with the overall service quality supplied at station. As for the attributes of service quality, users were asked to express the overall satisfaction giving a rate on an ordinal scale from 1 to 10.

The independent or exploratory variables are the selected attributes of service quality referred to the station and analysed in the present work.

The results are reported in Table 5. The coefficients reflect the strength and the type of relationship that the independent variables have to the dependent variable. We have hypothesized that each attribute of transit service quality has a positive impact on the overall service quality. Based on this assumption, the coefficients of the independent variables should have positive sign. This aspect is verified for all the variables; therefore, these variables are significant in the model.

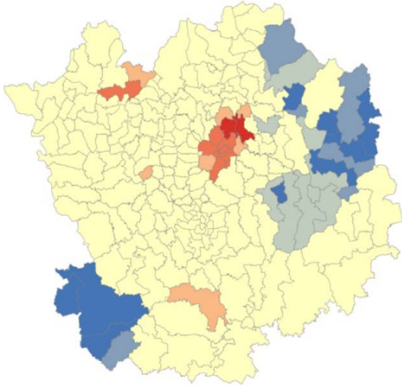
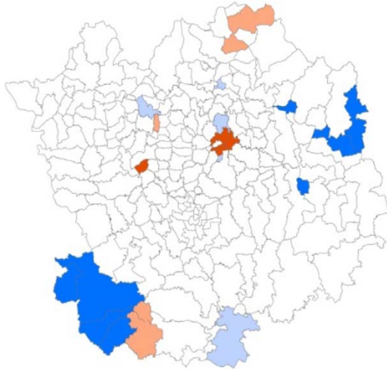
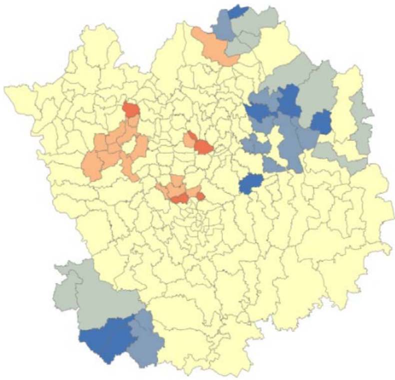
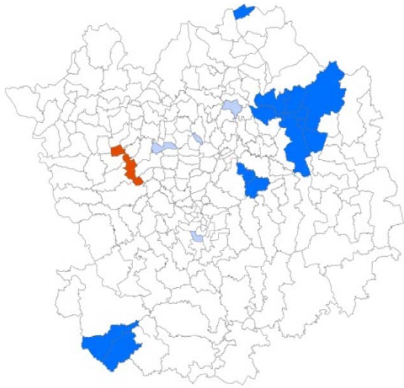
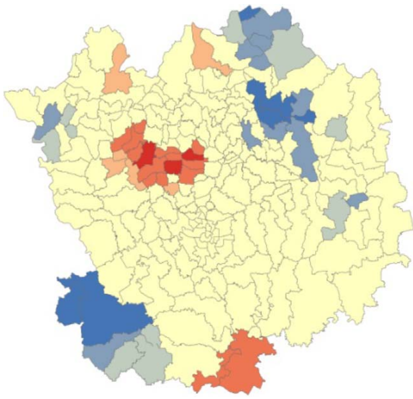
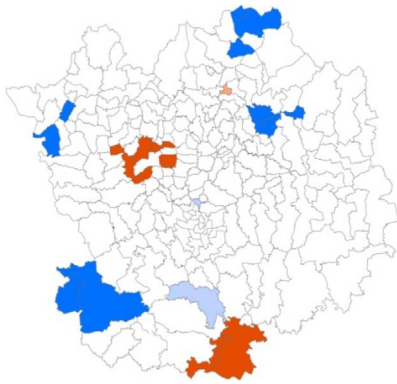
"Information at station" has the greatest impact on "Overall transit service quality", whereas "Facilities for disabled" influences in the way least significant the overall judgement on service quality, surely because it is an aspect which does not concern all the passengers.

Assessing the results of *t*-statistic, all the independent variables, including the intercept, are significant at 0.05 level. This means that the selected variables significantly impact on the overall satisfaction. The values of Variance Inflation Factor (VIF) indicate that all the independent variables are not redundant.

Regarding the goodness-of-fit of the model, R^2 and adjusted R^2, which give a measure of model performance, indicate that the model explains approximately 22% of the variation of the dependent variable. Without any doubt, the expected low value of the explanatory power of the model is due to the limited number of independent variables included in the model formulation. In fact, these variables relate only to certain attributes linked to the railway stations instead of considering the most appropriate attributes linked to service quality on board (as service frequency, service punctuality, cleanliness on board, etc.), which were surely considered by the users in their judgment when they expressed the satisfaction rate with the overall service.

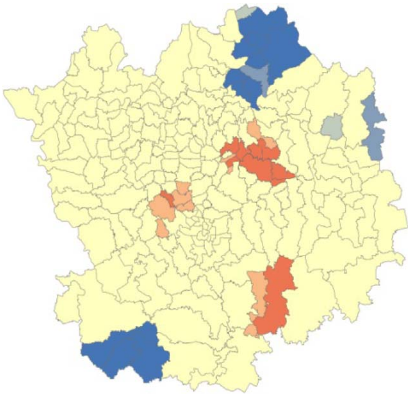
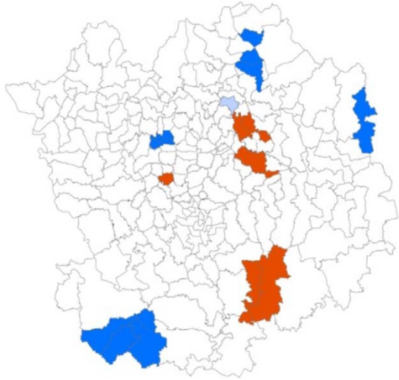
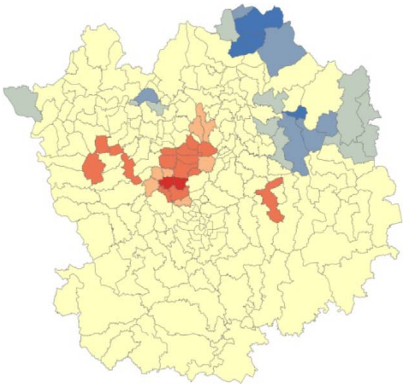
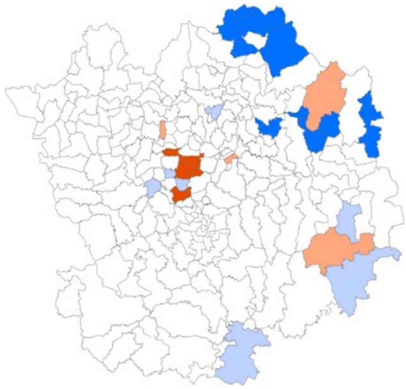
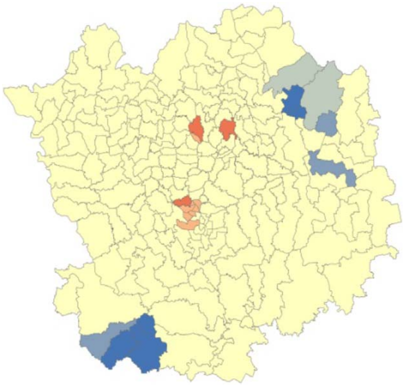
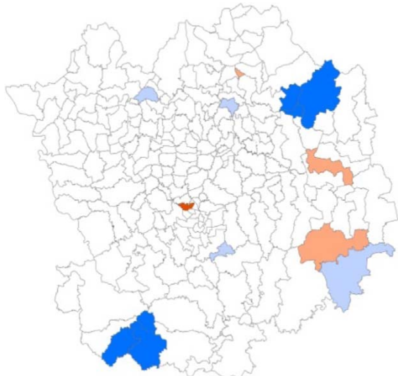
Akaike's Information Criterion (AIC) is a relative measure of performance used for comparing different models, as OLS and GWR. The model having the lowest AIC is the model that better explicates the dependent variable. F-Statistic, that is used for assessing overall model significance, results significant. Wald Statistic is used to assess overall robust model significance. In this case, the significant p-value indicated robust overall model significance. Koenker BP Statistic (Koenker's studentized Bruesch-Pagan statistic) is used to assess stationarity. The model is stationary when the it is consistent in geographic space, and the explanatory variables behave the same everywhere in the study area. The results obtained from this test were statistically significant and the null hypothesis that the model is stationary could be rejected.

Table 4
Results of local spatial autocorrelation tests.

Variable	Hot-spot analysis Getis-Ord	Anselin Local Moran's *I*
Personal security at station		
Cleanliness of station		
Maintenance of station		

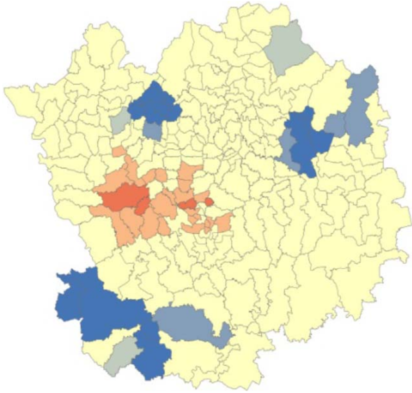
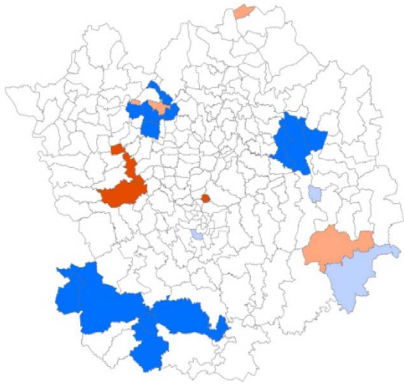
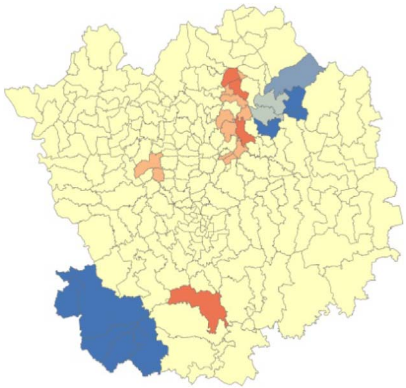
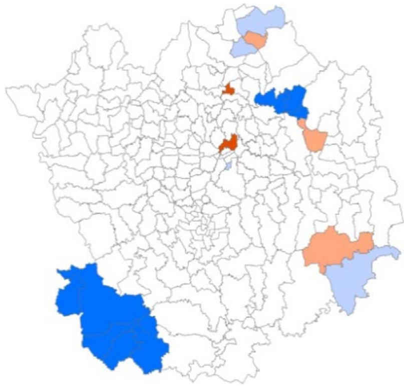
(*continued on next page*)

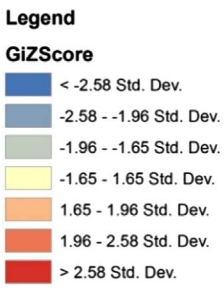
Table 4 (*continued*)

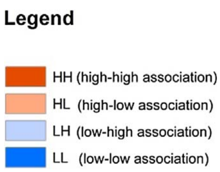
Variable	Hot-spot analysis Getis-Ord	Anselin Local Moran's *I*
Integration with other transit systems		
Facilities for disabled		
Information at station		

(*continued on next page*)

Table 4 (*continued*)

Variable	Hot-spot analysis Getis-Ord	Anselin Local Moran's *I*
Info connections with other transit systems		
Overall transit service quality		





The Jarque-Bera statistic indicates whether or not the residuals are normally distributed (the null hypothesis). In this case, the p-value for this test is smaller than 0.05 and the residuals are not normally distributed, indicating model misspecification.

As elaborated, the model presents no-stationarity and misspecification. The last step is assessing residual spatial autocorrelation, by means of performing the Moran's Index on residuals. In this case, residual errors showed spatial association. Regression models with statistically significant non stationarity are good candidates for GWR analysis.

GWR is based on the concept of distance decay, where more weight is given to the closer observations than the farther ones (Brunsdon et al., 1996). These weights are generated by means of a kernel function, which uses a bandwidth found by optimizing a goodness-of-fit criterion. The bandwidth can be considered the most important parameter for GWR because it controls the degree of smoothing in the model. Typically, a bandwidth or neighbour value can be obtained by selecting either Akaike Information Criterion (AICc) or Cross Validation (CV) for the bandwidth method parameter. Both of these options try to identify an optimal fixed distance or optimal adaptive number of neighbours (Eboli et al., 2013).

The proposed GWR model has the same variables that compare in the OLS and takes the following expression:

$$Y_i = \beta_0(u_i,v_i) + \sum_{k=1}^{7} \beta_k(u_i,v_i)X_{ki} \tag{1}$$

where

Table 5
Ordinary Least Squares results.

Variable	Coefficient	Std error	*t*-Statistic	Probability	VIF
Intercept	2.2406	0.0410	54.5337	0.0000*	–
Personal security at station	0.0967	0.0058	16.5276	0.0000*	1.3123
Cleanliness of station	0.1363	0.0096	14.1560	0.0000*	2.7564
Maintenance of station	0.0881	0.0095	9.2142	0.0000*	2.7601
Integration with other transit systems	0.0443	0.0047	9.4125	0.0000*	1.2285
Facilities for disabled	0.0146	0.0046	3.1364	0.0017*	1.1560
Information at station	0.1849	0.0062	29.9094	0.0000*	1.3824
Info connections with other transit systems	0.0507	0.0053	9.5164	0.0000*	1.3924
Number of observations	23,015		Number of variables		8
R^2	0.2210		Adjusted R^2		0.2207
AIC	966,698		Degrees of freedom		23,007
F-statistic	932.2314	Prob(> F), (7.230) degrees of freedom:			0.000000*
Wald statistic	5498.8838	Prob(> Chi-squared), (7) degrees of freedom:			0.000000*
Koenker (BP) statistic	178.1056	Prob(> Chi-squared), (7) degrees of freedom:			0.000000*
Jarque-Bera statistic	4719.9110	Prob(> Chi-squared), (2) degrees of freedom:			0.000000*

* Statistically significant at the 0.05 level.

- $\beta_0(u_i, v_i)$ is the intercept referred to the area i,
- $\beta_1(u_i, v_i)$, $\beta_2(u_i, v_i)$, $\beta_3(u_i, v_i)$, $\beta_4(u_i, v_i)$, $\beta_5(u_i, v_i)$, $\beta_6(u_i, v_i)$, and $\beta_7(u_i, v_i)$ are the coefficients of the explanatory variables referred to the area i;
- X_{1i}, X_{2i}, X_{3i}, X_{4i}, X_{5i}, X_{6i}, and X_{7i} are the independent variables referred to the area i.

The difference between OLS and GWR is that the coefficients β_k are constant in the global model, whereas in the local model they are unspecified functions of geographical coordinates (Brunsdon et al., 1999).

Table 6 shows the results of the statistics of goodness-of-fit regarding the overall model, which are useful in order to compare GWR with OLS. The bandwith was measured in meters. The matrix of spatial weights includes the distances calculated by considering the inverse Euclidean distance of the spatial units, and kernel is fixed. We adopted GWR because OLS is not suitable being the residuals of regression autocorrelated. The value of AIC is lower than in the case of OLS, therefore moving from a global model to a local regression model represents a good choice. In addition, R^2 and adjusted R^2 computed for GWR are higher than for OLS, and the amount of the variation in the dependent variable explained by the model GWR is approximately 25% and 23%, respectively. These values appear lower if compared with reference values. We have really expected such a result, because we did not use all the 33 service aspects (as independent variables) that influence the quality of the overall service (dependent variable). As specified in the paper, we considered only the service aspects concerning the station, that are correlated to the station. In this case, we consider them acceptable because in the analysis we put in relation with the satisfaction with the overall service quality only seven attributes. Consequently, the model does not explain all the variation in the dependent variable, and we waited a low value of R^2.

Although the explanatory power of the model is not high, the significance of the variables and the magnitude of their effects across modes are worth to examine. As checked by conditional numbers, local collinearity does not exist among the variables.

For understanding the influence of the explanatory variables on the dependent variable, the coefficient raster surfaces can be created and analysed. A coefficient raster surface is the graphic display of the coefficient of a certain variable. The analysis allows examining how spatially consistent relationships between dependent variable and each explanatory variable are across the study area. In OLS a unique value for each coefficient is related to the entire study area, without considering the spatial variation of variable across the areas. In addition, examining the coefficient distribution as a surface shows where and how much variation is present.

Fig. 3 shows the raster surfaces displaying the coefficient values obtained for the variables included in the model. We want to specify that the values of the coefficients represented in the figure are dimensionless and comparable among them because the users expressed own judgements about the aspects of transit service quality on the same scale of measurements. For each illustration, the values were divided into three classes, but being the maximum value of each coefficient different (while the minimum value is 0 for all the coefficients), the numerousness of each class varies among the coefficients.

Table 6
GWR diagnostic.

GWR Diagnostic	
Bandwidth	8679.92
Effective Number	394.53
Sigma	1.9610
AICc	96,544
R^2	0.2465
Adjusted R^2	0.2334

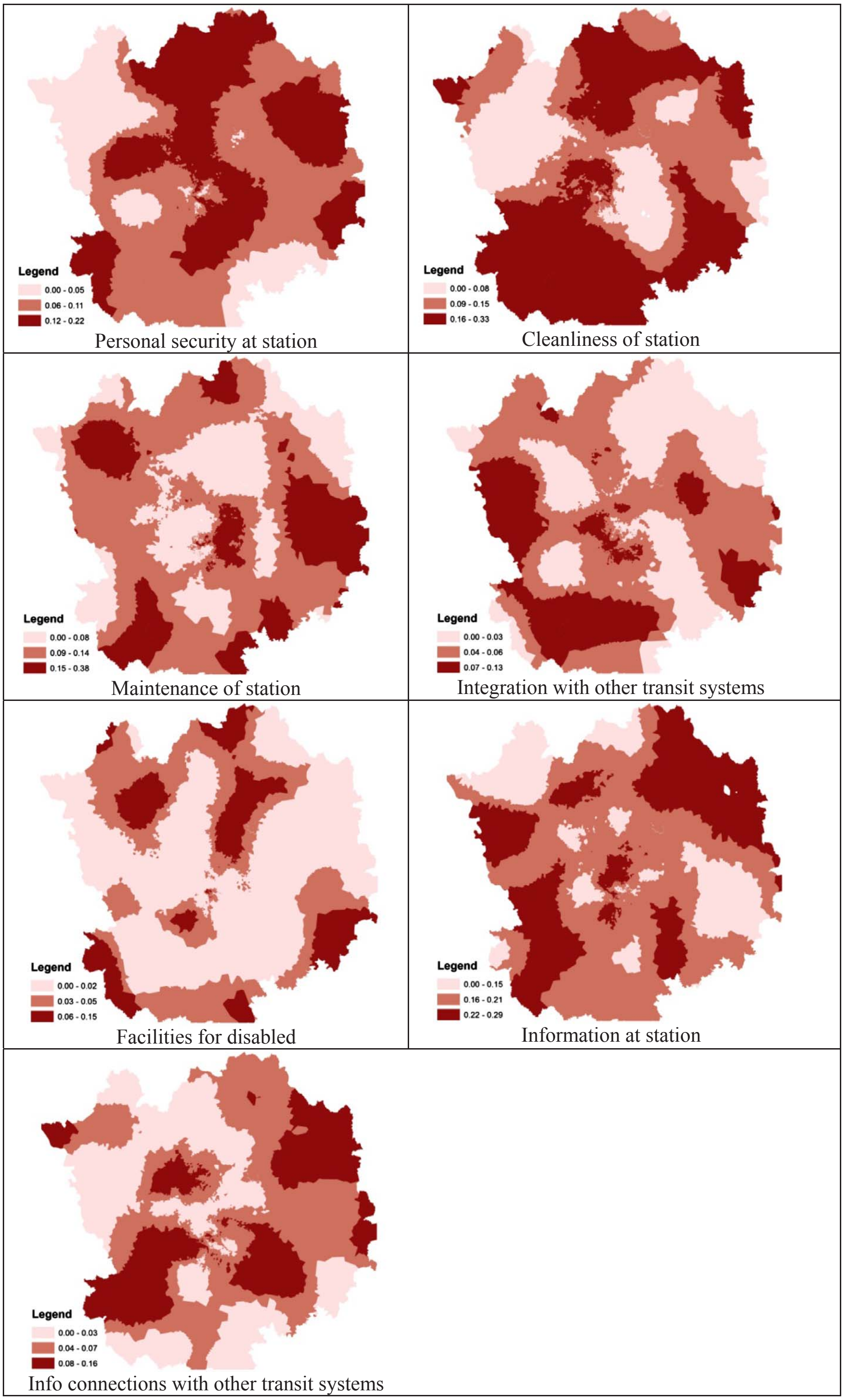


Fig. 3. Coefficient raster surfaces.

By observing the specific values of the classes in Fig. 3, we can say that there is a relevant variation among the three classes of users' satisfaction about each service aspect. As an example, by considering the service aspect concerning personal security at station, the average value of the first class is 0.025; the average value of the second class is 0.085, which is 3.4 times the average value of the first class; finally, the average value of the third class is 0.17, which is almost 7 times the average value of the first class. This means that there are parts of the study area where passengers are more sensitive to their personal security, probably because certain zones are less safe from this point of view. Similar observations and interpretation can be effected for the other service attributes. As an example, we can say that in some parts of the study area passengers are more sensitive to service aspects such as cleanliness or maintenance of station, because not all the stations offer to them satisfying levels of quality; surely, passengers perceive the degradation of some stations as regards others.

"Maintenance of station" is the attribute with the highest impact on the satisfaction with the overall service quality, whereas "Integration with other transit systems" has the lowest weight on the dependent variable, meaning that passengers retain most important to wait the train in a place characterized by a good level of quality, than having other transit means available for reaching the station. This last evidence is maybe due to the fact that the major part of the passengers have the possibility to reach the station by a own means of transport (car, foot, etc.).

By considering the raster surfaces, we can definitively conclude that service quality attributes have impact weights on the overall satisfaction rate which differ from a certain geographical area to another. As an example, in the major part of the cases we can note that in the central study area the weights of the attributes are lower than in the marginal areas. Specifically, certain areas (e.g. North and northeast areas, corresponding to Varese, Como and Lecco municipalities) present higher values of the coefficients except for "Integration with other transit systems" and "Facilities for disabled" attributes. In a general manner, we can observe the same tendency for the southwest area (Vigevano municipalities).

In any case, the application of the GWR methodology allows having a more appropriate interpretation of the impacts that each attribute can have on the overall customer satisfaction as a function of the geographical areas where passengers access to transit services.

5. Discussion and conclusion

In this study we have suggested a methodology for evaluating passengers' transit service quality satisfaction and treating spatial variation of service quality attributes across the study area. Particularly, we have focused on transit service quality evaluated at railway stations. The research was based on the data collected by means of a survey conducted by the railway service operating in the hinterland of Milan. Users were asked to express importance and satisfaction rates about service quality factors revealed on board and at the station where the trip starts. Among them, seven attributes concerning the service quality at station were selected, elaborated and included in the analysis.

The main objective of the paper was analysing the spatial variation of transit service quality evaluated at station. For this aim, we have elaborated a spatial regression named geographically weighted regression (GWR) in order to quantify how the influence of each attribute on the satisfaction with the overall service quality varies across the study area.

Firstly, we have applied techniques of spatial statistics in order to establish if the spatial pattern of data is the result of random choice or of spatial association. Global and local measures of spatial autocorrelation were applied on the attributes of transit service quality and implemented in a GIS environment. The outcomes of these spatial statistics have demonstrated that the values distribution of the attributes of service quality present spatial autocorrelation. As a consequence, no-spatial models seem to be inadequate and spatial regression model, as GWR, could be applied.

Compared to OLS, GWR has provided more appropriate results in terms of goodness-of-fit and significance of the model. Particularly, this analysis has allowed to examine the spatial relationships between the satisfaction with the overall service quality (dependent variable) and each attribute of transit service quality (explanatory variables) across the study area. By means of coefficient raster surfaces, the spatial variation has been graphically represented. We have been able to observe that the coefficients show a considerable spatial variation across the area, and at each station an attribute of service quality can have more influence than another on the overall users' satisfaction. This aspect could represent an innovative point of view for analysing transit service quality. In fact, the spatial analysis allows to locally studying users' satisfaction about transit service quality, instead of treating all the study area as a single part overlooking the spatial variability.

By applying spatial analysis techniques to the study area certain critical evidences were shown. Specifically, it emerges that the most critical areas are located in the southwest and northeast of the urban area of Milan. As expected, these areas are the most distant from Milan (more than 40 km); in addition, as we can note from Fig. 1b, in these areas there is a less dense railway network and the stations are more distant among them. On the contrary, the areas where railway service quality is perceived as higher are located in the Northern hinterland, near the city. These areas are characterized by a more intensive railway network, with frequent rail stations along the tracks.

From regression analysis we can conclude also that the above mentioned critical areas are characterized by passengers who are more sensitive to certain service attributes related to railway stations, because raster surfaces show average values of the coefficients higher than other areas. We specifically refer to cleanliness and maintenance of the stations, and also to information systems, by highlighting a general inaccuracy of the transit operator and local managements towards the seemliness of the stations in these peripheral areas.

Definitively, we can conclude that in the areas where the quality of service lack due to inappropriateness of the infrastructures offered by the land systems, passengers could become more sensitive in terms of perceptions about service quality aspects related to railway stations.

The empirical evidences derived from experimental data support our research hypothesis about the need of a methodology for evaluating passengers' satisfaction and treating spatial variation of service attributes across the considered study area.

A deep spatial knowledge of attributes of transit service quality could led to a better use of economic and energy resources in order to provide an efficient transit service. In fact, targeted interventions on certain aspects and specific areas can reduce the investment needed to get improvements in the service.

We really believe that our work provide a useful contribution in terms of new knowledge, because we apply techniques of spatial analysis in the field of customer satisfaction and transit service quality measure, which was traditionally studied without taking into account the variation of the quality of some service quality attributes in the space.

As above specified, the overall service quality has been traditionally measured as a result of the service quality of the single service aspects. There are very many techniques proposed in the literature adopted for explaining how overall service quality depends on the quality of the various aspects describing the service. The experts in the sector of service quality and customer satisfaction know very well this issue. In this paper, we try to give a more innovative contribution to the literature because we investigate the issue by taking into account the space variable.

Acknowledgements

Support from Spanish Ministry of Economy and Competitiveness (Research Project TRA2015-66235-R) is gratefully acknowledged.

References

Abenoza, R.F., Cats, O., Susilo, Y.O., 2017. Travel satisfaction with public transport: determinants, user classes, regional disparities and their evolution. Transp. Res. Part A: Policy Pract. 95 (1), 64–84.

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.

Anselin, L., 1995. Local Indicators of Spatial Association-LISA. Research Paper 9331, Regional Research Institute West Virginia University Morgantown.

Anselin, L., Sridharan, S., Gholston, S., 2007. Using exploratory spatial data analysis to leverage social indicators databases: the discovery of interesting patterns. Soc. Indic. Res. 82, 287–309.

Ayyildiz, K., Cavallaro, F., Nocera, S., Willenbrock, R., 2017. Reducing fuel consumption and carbon emissions through eco-drive training. Transp. Res. Part F: Psychol. Behav. 46, 96–110. http://dx.doi.org/10.1016/j.trf.2017.01.006.

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

Brunsdon, C., Fotheringham, A.S., Charlton, M., 1996. Geographically weighted regression: A method for exploring spatial non-stationarity. Geogr. Anal. 28 (4), 281–298.

Brunsdon, C., Fotheringham, A.S., Charlton, M., 1999. Some notes on parametric significance test for geographically weighted regression. J. Regional Sci. 39 (3), 497–584.

Cardozo, O.D., García-Palomares, J.C., Gutiérrez, J., 2012. Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. Appl. Geogr. 34, 548–558.

Cats, O., Abenoza, R., Liu, C., Susilo, Y., 2015. Identifying priority areas based on a thirteen years evolution of satisfaction with public transport and its determinants. Transp. Res. Rec. J. Transp. Res. Board 2323, 99–109.

Cavana, R.Y., Corbett, L.M., Lo, Y.L., 2007. Developing zones of tolerance for managing passenger rail service quality. Int. J. Qual. Reliab. Manage. 24 (1), 7–31. http://dx.doi.org/10.1108/02656710710720303.

Clark, S.D., 2007. Estimating local car ownership models. J. Transp. Geogr. 15, 184–197.

Cirillo, C., Eboli, L., Mazzulla, G., 2011. On the asymmetric user perception of transit service quality. Int. J. Sustain. Transp. 5 (4), 216–232. http://dx.doi.org/10.1080/15568318.2010.494231.

Comité Européen de Normalisation (CEN), 2002. European Standard EN 13816. Transportation, Logistics and Service, Public Passenger Transport, Service Quality Definition, Targeting and Measurement.

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

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.

de Oña, J., de Oña, R., Eboli, L., Mazzulla, G., 2016. Index numbers for monitoring transit service quality. Transp. Res. Part A 84, 18–30. http://dx.doi.org/10.1016/j.tra.2015.05.018.

de Smith, M.J., Goodchild, M.F., Longley, P.A., 2009. Geospatial Analysis. A Comprehensive Guide to Principles, Techniques and Software Tools. Matador, Leicester.

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

Diab, E., El-Geneidy, A., 2012. Understanding the impacts of a combination of service improvement strategies on bus running time and passenger's perception. Transp. Res. Part A: Policy Pract. 46 (3), 614–625. http://dx.doi.org/10.1016/j.tra.2011.11.013.

Diab, E., El-Geneidy, A., 2014. Transitory optimism: changes in passenger perception following bus service improvement over time. Transp. Res. Rec. J. Transp. Res. Board 2415, 90–101. http://dx.doi.org/10.3141/2415-11.

Eboli, L., Forciniti, C., Mazzulla, G., 2013. Exploring the relationship among urban system characteristics and trips generation through a GWR. Int. J. Innov. Inf. Technol. 1 (2), 51–61.

Eboli, L., Forciniti, C., Mazzulla, G., Calvo, F., 2016. Exploring the factors that impact on transit use through an ordered probit model: the case of metro of Madrid. Transp. Res. Procedia 18, 35–43. http://dx.doi.org/10.1016/j.trpro.2016.12.005.

Eboli, L., Mazzulla, G., 2008a. A stated preference experiment for measuring service quality in public transport. Transp. Plann. Technol. 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. Eur. Transp. – Trasporti Europei 38, 107–118.

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

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. Transp. Policy 18 (1), 172–181. http://dx.doi.org/10.1016/j.tranpol.2010.07.007.

Eboli, L., Mazzulla, G., 2012a. Performance indicators for an objective measure of public transport service quality. Eur. Transp. – Trasporti Europei 51, 1–21.

Eboli, L., Mazzulla, G., 2012b. Structural equation modelling for analysing passengers' perceptions about railway services. Procedia – Soc. Behav. Sci. 54, 96–106. http://dx.doi.org/10.1016/j.sbspro.2012.09.729.

Fotheringham, A.S., Brunsdon, C., Charlton, M.E., 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Wiley, Chichester.

Getis, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. Geogr. Anal. 24 (3), 189–206. http://dx.doi.org/10.1111/j.1538-4632.1992.

tb00261.x.
Griffith, D.A., Layne, L.J., 1999. A Casebook for Spatial Statistical Data Analysis: A Compilation of Analyses of Different Thematic Data Sets. Oxford University Press, New York.
Grisé, E., El-Geneidy, A., 2017. Evaluating the relationship between socially (dis)advantaged neighborhoods and customer satisfaction of bus service in London, U.K. Transp. Geogr. 58, 166–175. http://dx.doi.org/10.1016/j.jtrangeo.2016.11.016.
Habib, K.M.N., Kattan, L., Islam, T., 2011. Model of personal attitudes towards transit service quality. J. Adv. Transp. 45 (4), 271–285. http://dx.doi.org/10.1002/atr.106.
Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010. Development of planning level transportation safety tools using geographically weighted poisson regression. Accid. Anal. Prev. 42, 676–688. http://dx.doi.org/10.1016/j.aap.2009.10.016.
Hensher, D.A., Prioni, P., 2002. A service quality index for area-wide contract performance assessment. J. Transp. Econ. 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. Transp. Res. Part A: Policy Pract. 37 (6), 499–517. http://dx.doi.org/10.1016/S0965-8564(02)00075-7.
Hernandez, S., Monzon, A., de Oña, R., 2016. Urban transport interchanges: A methodology for evaluating perceived quality. Transp. Res. Part A: Policy Pract. 84, 31–43. http://dx.doi.org/10.1016/j.tra.2015.08.008.
Iseki, H., Taylor, B.D., 2010. Style versus service? An analysis of user perceptions of transit stops and stations. J. Public Transp. 13 (3), 39–63.
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. J. Public Transp. 10 (4), 99–118.
Lao, Y., Liu, L., 2009. Performance evaluation of bus lines with data envelopment analysis and geographic information systems. Comput. Environ. Urban Syst. 33 (4), 247–255. http://dx.doi.org/10.1016/j.compenvurbsys.2009.01.005.
Lloyd, C.D., 2010. Local Models for Spatial Analysis. Taylor & Francis, Boca Raton.
Nathanail, E., 2008. Measuring the quality of service for passengers on the Hellenic railways. Transp. Res. Part A: Policy Pract. 42 (1), 48–66. http://dx.doi.org/10.1016/j.tra.2007.06.006.
Nkurunziza, A., Zuidgeest, M., Brussel, M., Van den Bosch, F., 2012. Spatial variation of transit service quality preferences in Dar-es-Salaam. J. Transp. Geogr. 24, 12–21. http://dx.doi.org/10.1016/j.jtrangeo.2012.06.001.
Nocera, S., Cavallaro, F., 2016. The competitiveness of alternative transport fuels for CO_2 emissions. Transp. Policy 50, 1–14. http://dx.doi.org/10.1016/j.tranpol.2016.05.013.
Ord, J.K., Getis, A., 1995. Local spatial autocorrelation statistics: distributional issues and an application. Geogr. Anal. 27, 286–306.
Ord, J.K., Getis, A., 2001. Testing for local spatial autocorrelation in the presence of global autocorrelation. J. Regional Sci. 41 (3), 411–432.
Páez, A., 2006. Exploring contextual variations in land use transport analysis using a probit model with geographical weights. J. Transp. Geogr. 14 (3), 67–176.
Pakdil, F., Aydın, O., 2007. Expectations and perceptions in airline services: an analysis using weighted SERVQUAL scores. J. Air Transp. Manage. 13 (4), 229–237. http://dx.doi.org/10.1016/j.jairtraman.2007.04.001.
Sheth, C., Triantis, K., Teodorović, D., 2007. Performance evaluation of bus routes: a provider and passenger perspective. Transp. Res. Part E: Logist. Transp. Rev. 43 (4), 453–478. http://dx.doi.org/10.1016/j.tre.2005.09.010.
Weinstein, A., 2000. Customer satisfaction among transit riders: how customers rank the relative importance of various service attributes. Transp. Res. Rec. 1735, 123–132. http://dx.doi.org/10.3141/1735-15.