

EXPLORING ON-DEMAND SERVICE USE IN LARGE URBAN AREAS: THE CASE OF ROME

Agostino NUZZOLO¹, Antonio COMI², Antonio POLIMENI³

^{1,2,3} University of Rome Tor Vergata, Department of Enterprise Engineering, Rome, Italy

Abstract:

Traditional and innovative on-demand transport services, such as taxi, car sharing or dial-a-ride respectively, can provide a level of flexibility to the public transport with the aim to guarantee a better service and to reduce the exploitation costs. In this context, in order to point out the key-factors of on-demand services, this study focuses on traditional on-demand service (such as taxi one), and presents the results of a demand analysis and modelling, obtained processing taxi floating car data (FCD) available for the city of Rome. The GPS position of each taxi is logged every few seconds and it was possible to build a monthly database of historical GPS traces through around 27 thousands of GPS positions recorded per day (more than 750 thousands for the entire month). Further, the patterns of within-day and day-to-day service demand are investigated, considering the origin, the destination and other characteristics of the trips (e.g. travel time).

The time-based requests for taxi service were obtained and used to analyse the trip distribution in space and on time. These analyses allow us to forecast trips generated/attracted by each zone within the cities according to land use characteristics and time slices. Therefore, a regression tree analysis was developed and different regressive model specifications with different set of attributes (e.g. number of subway stations, number of zonal employees, population) were tested in order to assess their contribution on describing such a service use.

Keywords: on-demand services, demand analysis, decision tree.

To cite this article:

Nuzzolo, A., Comi, A., Polimeni, A., 2019. Exploring on-demand service use in large urban areas: the case of Rome. *Archives of Transport*, 50(2), 77-90. DOI: <https://doi.org/10.5604/01.3001.0013.5681>



Contact:

1) nuzzolo@ing.uniroma2.it [<http://orcid.org/0000-0002-9685-4849>], 2) comi@ing.uniroma2.it [<http://orcid.org/0000-0001-6784-1638>], 3) antonio.polimeni@uniroma2.it [<http://orcid.org/0000-0002-5104-7373>] – corresponding author

1. Introduction

Inner urban areas are the ones that suffer more than others the negative impacts on high volume of private traffic, such as, among others, congestion, air pollution, and health risks. Improvement of air quality and reduction of congestion in central areas has been a priority in many cities (e.g. Paris, London, Rome; Russo and Comi, 2016; Comi et al., 2019; Font et al., 2019; Piras et al., 2019), and regulatory traffic control (Alonso et al., 2019), usually integrated with the improvement of public transport, has been implemented (Suman et al., 2019). In this context, on-demand services (e.g. taxi, car sharing or dial-a-ride) significantly can contribute to reduce traffic congestion, energy consumption, and air pollution thanks the opportunity offered in improving the vehicle use and the optimization of vehicles-kms. Understanding the travel patterns of such services is thus important for addressing many urban sustainability challenges (SUMP, 2013). Besides, having a better understanding of the passenger demand over different spatial zones is of great importance to better organize service locations and, hence, to improve the vehicle utilization rate, limiting their impacts on urban environment.

Although limited research efforts have been implemented on forecasting such a demand, in most recent years mainly pushed by the real-world data unavailability, the studies on the taxi service (i.e. traditional on-demand service) can provide valuable insights since there exist strong similarities between the taxi and the on-demand services (Moreira-Matias et al. 2013 a and b; Nuzzolo et al., 2019). Besides, the analyses on taxi demand can be helpful to define the characteristics of future autonomous vehicles demand in mobility scenarios (Czech, 2018).

In this background, the first objective of paper germinates, i.e. to propose a methodology for investigating taxi demand included the test of its goodness. The case study refers taxi service in Rome, where taxis are traced, and their positions on time are provided. The first stage of the analysis is addressed to define the taxi status (i.e. driving with a customer, driving to a customer, at parking) analysing its movements in the space and on time. This allows us to define time-dependent origin-destination trips (with the relative travel time and distance attributes) when customers are on board providing information for developing methods and models for forecasting on-demand service use. Subsequently, pushed by the

opportunity to have models and methods for demand forecasting, the second objective of paper is to develop a set of regressive models, able to foresee the taxi user demand in specific time intervals. This can help taxi service to be improved and to provide tools for on-demand service request forecasting.

The paper is organized as follows. Section 2 reports a brief literature review. Section 3 reports a spatial and temporal analysis of data collected in the city of Rome. Section 4 deals with a descriptive analysis of the results, developing a regression tree and models. Some general conclusions and further research developments are considered in Section 5.

2. Literature review

The use of new technologies, like global positioning system - GPS, can help the analysis of vehicle movements obtaining large quantity of data continuously in space and on time at low costs. Usually, such data were used for planning and operations control of transit services (Nuzzolo and Comi, 2017 and references cited therein) and for commercial fleet management and control (Polimeni and Vitetta, 2014; Hadavi et al., 2018; Croce et al., 2019). Recently, the availability of these data also from taxi services has allowed us to point out this segment of demand, which is so significant in inner urban areas (Li et al., 2019).

Taxi data are usually used to evaluate the status of the taxi transport system in a city, ranging from the analysis of the taxi trips to the demand analysis. Bischoff et al. (2015) analysed the taxi service in Berlin considering the travel behaviour and vehicle supply using floating car data (FCD). Other examples of taxi data processing and interpretation are reported in Liu et al. (2015) that analysed the trips of about 6000 taxi to identify the spatial interactions between areas of the city and in Nuzzolo et al. (2019) that analysed the patterns of within-day and day-to-day service demand.

Taxi management and control activities are other aspects considered in literature. Wang et al. (2015) developed a dynamic model for urban taxi management, to simulate and to forecast the possible conditions of taxi operations. Main attributes considered in the model are empty-loaded rate, total demand, and the impacts of taxi fares. A model for supporting the decision and policy making in taxi market is proposed by Grau and Estrada (2019), considering some indicators as waiting/access time and benefits of the

drivers. Moreover, the model considers the demand elasticity, allowing the optimum fare for taxi services to be estimated. Zheng et al. (2018) proposed a model to simulate drivers' behaviour to anticipate the variation (in time) of demand at locations such as airports and train stations. Defined a decision horizon, the model simulates as taxi driver choices to transfer to such a destination (considering a reward function and the taxi driver learning from experiences). The uncertainty of demand is simulated by a pick-up probability. Xiao and He (2017) proposed a multi-objective model and a subsequent solution algorithm to optimize taxi carpooling. The aim is to minimize the access/egress distance and the carpooling waiting time.

Tang et al. (2015) and Moreira-Matias et al. (2016) proposed a methodology to build the origin-destination matrix. The taxi travels can be analysed to individuate the empty travels and the travel with a customer (Cai et al., 2016; Wang et al., 2015). These data (Yang et al., 2018) can be used to define a model linking demand for taxi with other city characteristics (densities, accessibility to other modes). He and Shen (2015) developed a spatial equilibrium model that tries to capture the taxi drivers and customer's adoption of the e-hailing applications. The model considers both taxi drivers and customer's flexibility in the choice of the e-hailing application, and vacant taxi drivers can determine where to look the customers. Al-Ayyash et al. (2016) proposed a model framework to evaluate the market demand potential of a shared ride taxi service; the goal is the simulation of the number of times per week the service would be used. A case study is proposed to investigate whether students would utilize the service if it were implemented. In relation to the taxi operation modelling, Wong et al. (2001) set up a two-level model: the upper level simulates the relation between taxi and customer waiting times and the relation between customer demand and taxi supply; the lower level problem is a combined network equilibrium model (that describes simultaneous movements of vacant and occupied taxis and of the normal traffic). An extension of this model, considering multiple classes of users and taxi services, is reported in Wong et al. (2008) where the mileage-based and congestion-based taxi fare charging mechanisms is modelled in a unified framework. A modal choice of customers for taxi services is also introduced. In relation to the topic of travel demand forecasting,

Moreira-Matias et al. (2013a) used time series approaches to foresee (short-term horizon) the spatial distribution of the users. An extension of this study is reported in Moreira-Matias et al. (2013b) where the demand forecasting is approached with a learning model to foresee the demand every 5 minutes. Ramezani and Nourinejad (2018) modelled taxi dispatching, linking the urban traffic flow with the taxi operations. The model is based on the macroscopic fundamental diagram (in this approach, the study area is partitioned into multiple zones, each one represented by a macroscopic fundamental diagram) that simulates the dynamic traffic conditions. The taxi-dispatching problem is tackled in Maciejewski (2014), where two different approaches are theorized to approximate the upper and lower bounds for the passenger waiting time.

In relation to taxi user perceived quality, Alonso et al. (2019) estimated two types of models: i) a model that simulates the perceived quality of taxi service when no previous information about the system is available and ii) a model to simulate the changes in users' perception after being informed about the attributes. The goal is to understand and improve the system in relation to the user preferences.

From the literature review synthesized above, it emerges that although there is a growing interest in proposing methods and models for simulating on-demand services, few of them investigated the main factors influencing demand and proposed forecasting models. This shows that further work needs in this field. Therefore, this paper, moving from a first study (Nuzzolo et al., 2019) where some statistical descriptive analysis are performed, presents the results of a methodology developed for analysing and modelling the spatial-temporal taxi demand and the results of the study case developed in Rome. Besides, the advancement proposed with respect the current state-of-the-art (Moreira-Matias et al., 2013a and b; Tang et al., 2015; Moreira-Matias et al., 2016) mainly refer to the development of models that use easy-to-obtain variables (such as data from national census, e.g. population, employees for type) and to the use of machine learning techniques for identifying significant predictors of regressive forecasting models.

3. Data analysis

The developed methodology provides to start from the current taxi service pattern to identify the drivers

and characteristics of demand on time and in space. The methodology was implemented in Rome where a large dataset containing the taxi data of a large company in the month of February 2014 (Bracciale et al., 2014) was available.

The data consists of about 756 thousand records, related to 312 taxi driver operations on about 7700 taxi in the entire city. Each data point contains a timestamp up to seconds of when the data was recorded, vehicle ID, vehicle location (i.e. longitude and latitude) at the recording time.

The identification of the study area and its zoning allows us to delimitate the geographical area that includes the transportation system under analysis and encompasses most of the project effects.

The dataset is hence checked for data quality and cleaned, e.g. some trips could conduct outside of the study area borders, therefore, all trips outside a given coordinate range are dropped following, for example, literature methods (e.g. Barann et al., 2017). Besides, the collected data do not contain explicitly the taxi status (e.g. empty or with customer). Then, a procedure was developed aiming to detect one of the following situations: driving with a customer, driving to a customer, at parking. The main interest was to find the travels with a customer on board to define a space-temporal structure of the travels.

The analysis of taxi travels from GPS data can be resumed as (fig. 1. shows the procedure for one taxi):

- *Vehicle status identification*; if the distance travelled in two minutes is less than ten meters, the taxi is considered to be non-moving, otherwise it is assumed traveling;
- *Origin-destination individuation*; for each travel, origin and destination are evaluated;
- *Travel distance*; for each travel, the travelled distance is evaluated from recorded positions;
- *Average travel time*; for each travel, the travel time is calculated considering the time from GPS data;
- *Waiting time*; it is the time spent by time at same position (which is also identified as the trip destination).

The taxi status *with customer* is the situation considered in this paper in order to individuate origin and destination of taxi trips. Since this condition is not explicitly provided by the available data, some assumptions are needed to individuate it. In order to individuate this status, it must be established when the vehicle is at a stand. A taxi is considered to be in a stand when two conditions happen:

- 1) the recorded vehicle position is close to a taxi stands in the city (in Rome, there are more than 100 taxi stands) and
- 2) the value of the waiting time is more than 2 minutes.

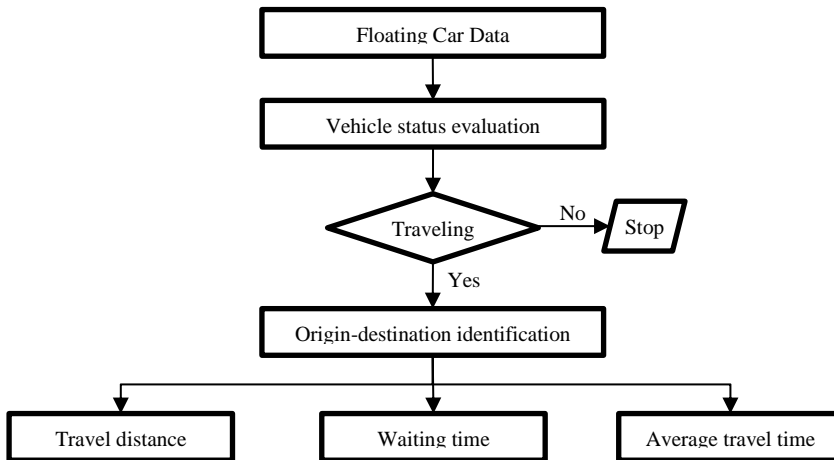


Fig. 1. Data analysis: procedure for taxi travel analysis

In this way, it is possible to evaluate the number of times when a vehicle is steady at a stand and when it is steady because the driver is dealing with a customer (the customer is entering or is leaving). Consequently, it can be found the trips with the activity *with customer*. Supposing that the first travel in a

day start from a stand (this assumption is consistent with the behaviour of Italian taxi drivers; Nuzzolo et al., 2019), with the previous assumptions it is possible to reconstruct the daily movements of a taxi, and then to calculate the average trip travel time (fig. 2.) and length (fig. 3.)

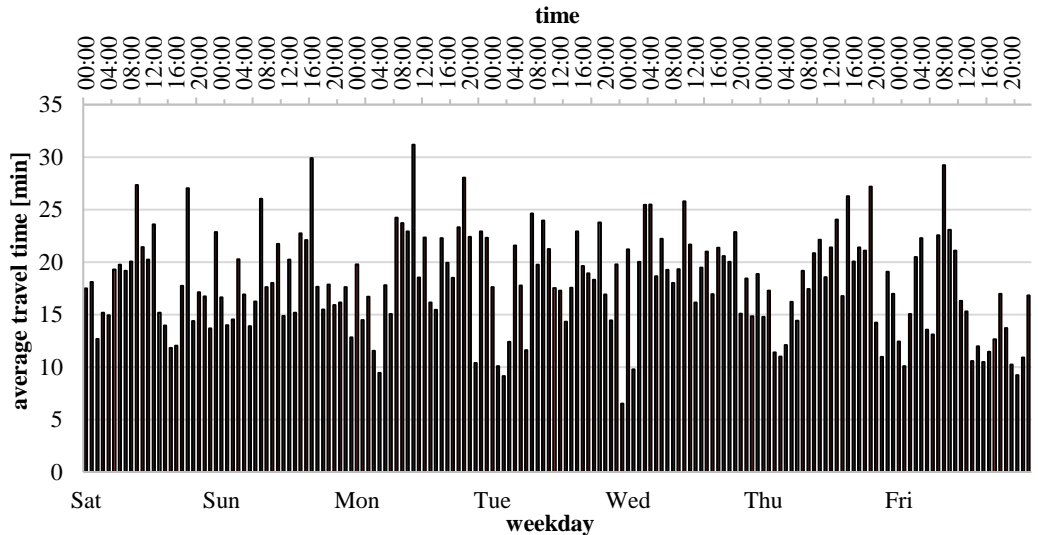


Fig. 2. Average trip travel time

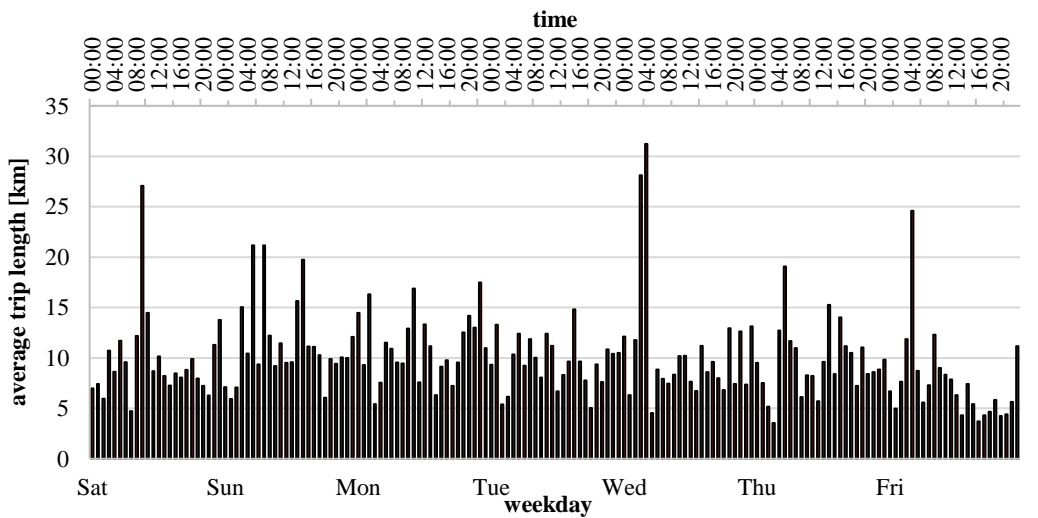


Fig. 3. Average trip length

A first analysis focused on the distribution of taxi demand (i.e. the number of requests submitted per day and per hour) in the first two weeks of February and showed the relevance to point out the spatial dimension given that significant differences exist in temporal patterns among different areas of the city. Besides, relevant differences were revealed on working days and weekend. Therefore, moving from these first results we explore the spatial and temporal characteristics of taxi demand for three spatial cases: the taxi requests in the zones of the main railway stations (i.e. Roma Termini station and Roma Tiburtina station) and the requests registered in the other zones. The analysis was carried out only for the weekdays (20 days).

3.1. Spatial analysis

The municipality of Rome extends over an area of 1,283.70 km², with the 22.20% destined to urban activities. According to the general urban traffic plan (Piano Generale del Traffico Urbano – PGTU, Roma Capitale, 2015), the area is divided in six macro-areas. Moreover (fig. 4.), each of them was partitioned into 99 homogeneous traffic zones (in turn, each

traffic zone is from the union of some Italian census areas). However, for the purpose of this paper, only the zones inside the main road ring (GRA) are considered, where most of travel taxi departures and arrivals were observed. The zones containing the railway stations are labelled with 4 (Roma Termini station) and 18 (Roma Tiburtina station).

Origins and destinations of taxi trips are generally spread all over the city. The majority starts and ends within the city centre: as an example, in fig. 5 the zones 1, 4 and 26 have a taxi request (related with the analysed sample) greater than 20%.

The data also show that there is a significant rate of trips from the city to the outside and vice versa; this is explainable with the fact that Fiumicino Airport, the biggest Rome Airport, is outside the city limits.

Points of special interest are often origin/destination for taxi trips. These include railway stations, the fair grounds and major event locations. Among all the railway stations with regular long-distance connections, the Roma Termini station has the biggest attraction (in terms of origin/destination) for taxi traffic, followed by the Tiburtina station.

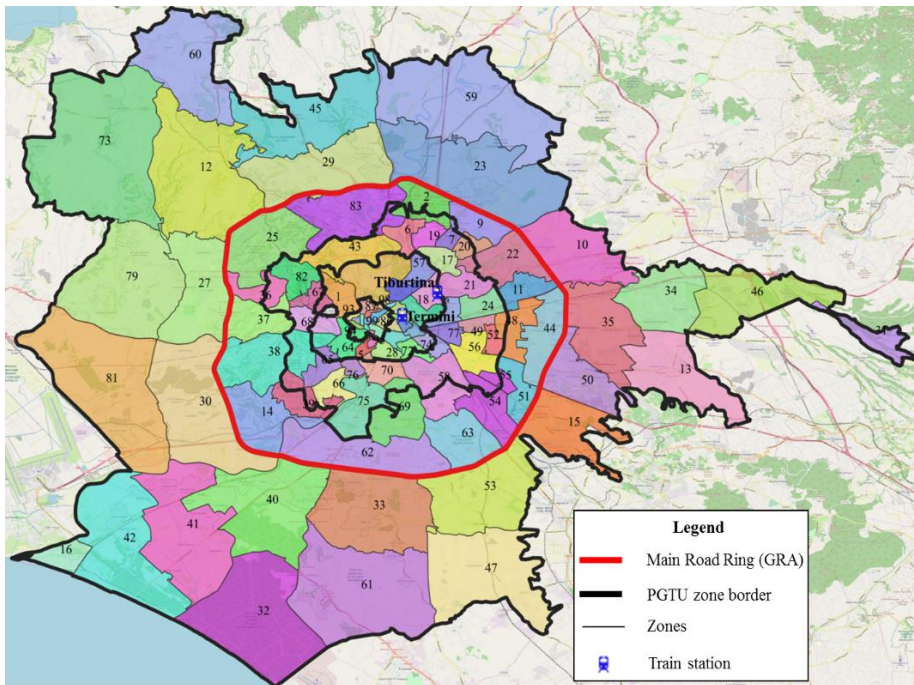


Fig. 4. Zoning

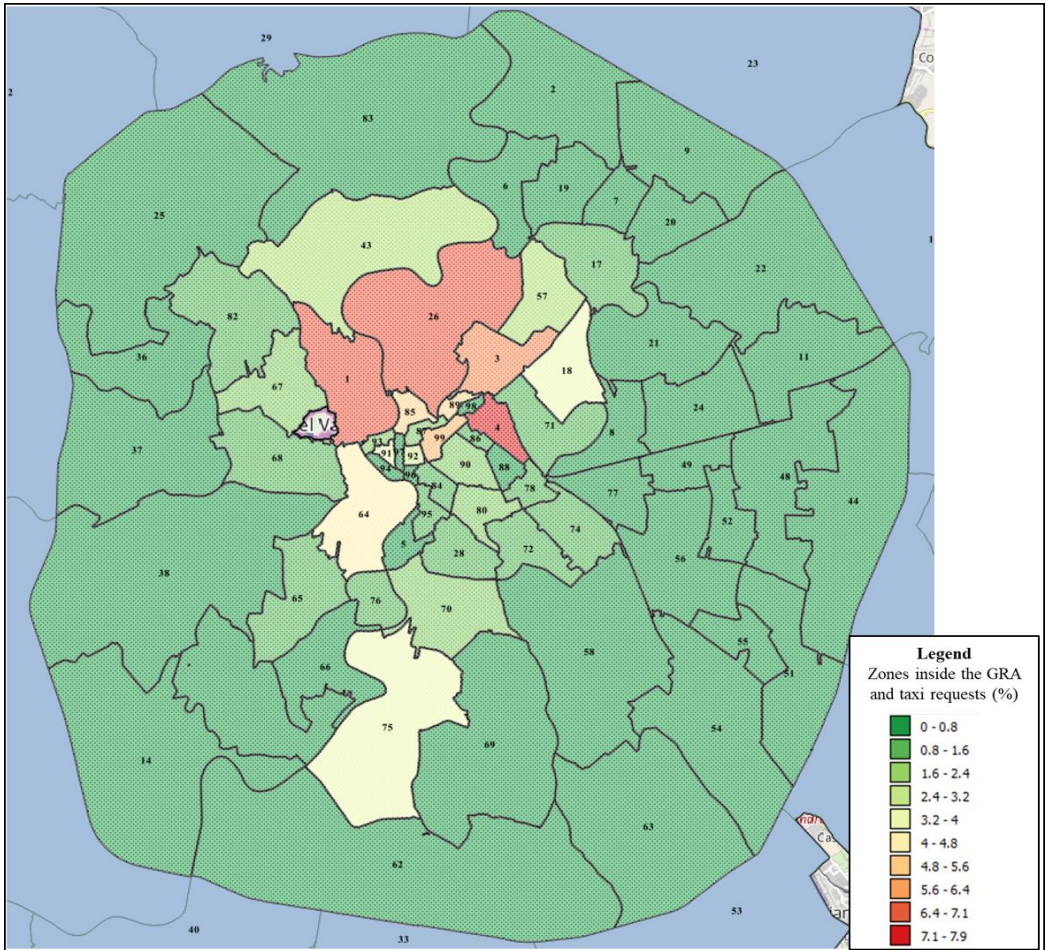


Fig. 5. Zones inside and outside the GRA (i.e. main road ring) and taxi requests

3.2. Temporal analysis

As aforementioned, the data are related to four weeks in the month of February 2018. Only the data related to working days were considered in this study. Subject to the considerations made in the previous section, the aggregated requests for weekday (considering the zones inside the study area without the requests from/to train stations) are represented in fig. 6. Similarly, fig. 7 reports the same data for the two train stations. It can be observed that in the analysed period (and in relation with the sample of available data) the taxi requests in the zones containing the train stations range from 7.08% (Tiburtina station) to 16.52% (Termini station). Besides, it

emerges that the requests related with the two train stations are about the 24% of the total taxi requests in the area (Table 1).

Table 1. Aggregated requests for day

Weekday	Termini	Tiburtina	Other	Total
Mon	17.80%	7.08%	75.12%	100%
Tue	16.52%	7.36%	76.12%	100%
Wed	15.61%	8.46%	75.93%	100%
Thu	14.02%	8.46%	77.53%	100%
Fri	15.01%	8.63%	76.36%	100%
Mean	15.79%	8.00%	76.21%	

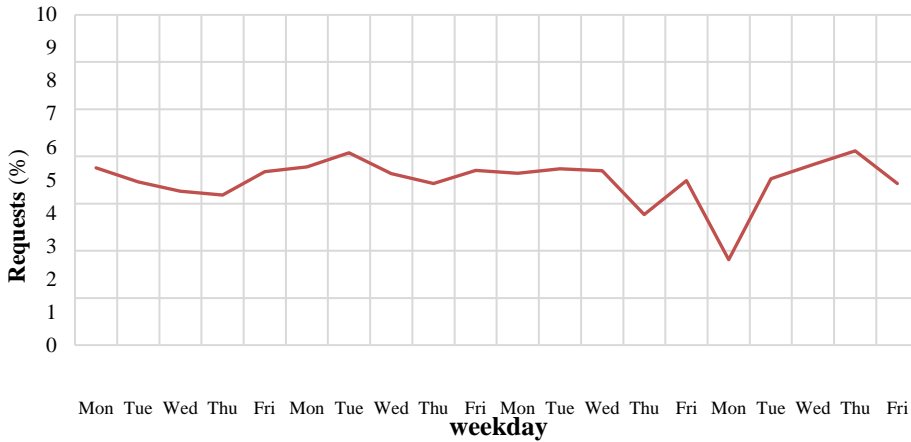


Fig. 6. The total requests for day (excluding the train stations)

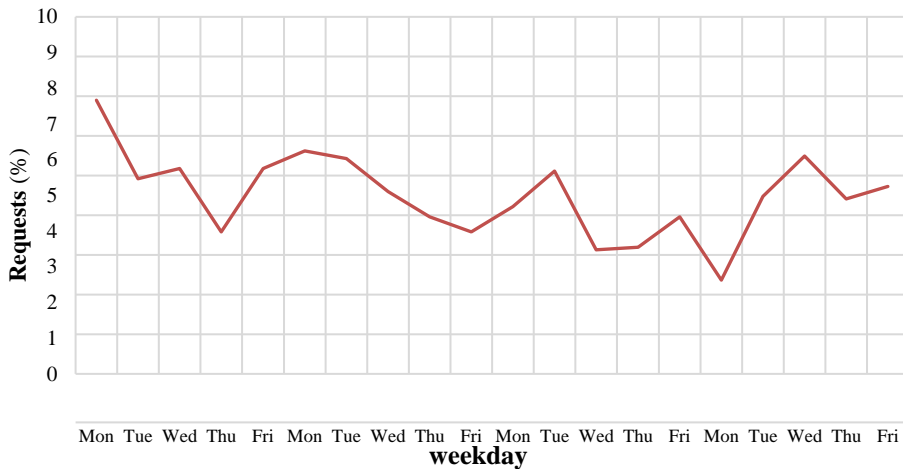


Fig. 7. The total requests for day from/to Termini train station

A more disaggregate analysis was performed, considering the temporal distribution of the taxi requests for each day per half an hour in the study area from/to Termini station (fig. 8.). The effect of time emerges showing that six main time slices can be identified:

- Time slice 1 (FA1): [00.00-04:30), nightly requests also related to the closure of metro services;
- Time slice 2 (FA2): [04:30-07:00), first daily requests for very early morning departures/arrivals;
- Time slice 3 (FA3): [07:00-10:00), requests for early morning departures/arrivals;
- Time slice 4 (FA4): [10:00-11:30), morning requests to and from city activities;
- Time slice 5 (FA5): [11:30-16:00), daily requests to and/or from transport hub, as well as from touristic places (including hotels);
- Time slice 6 (FA6): [16:00-00:00), daily requests to and/or from transport hub, as well as to touristic places (including hotels).

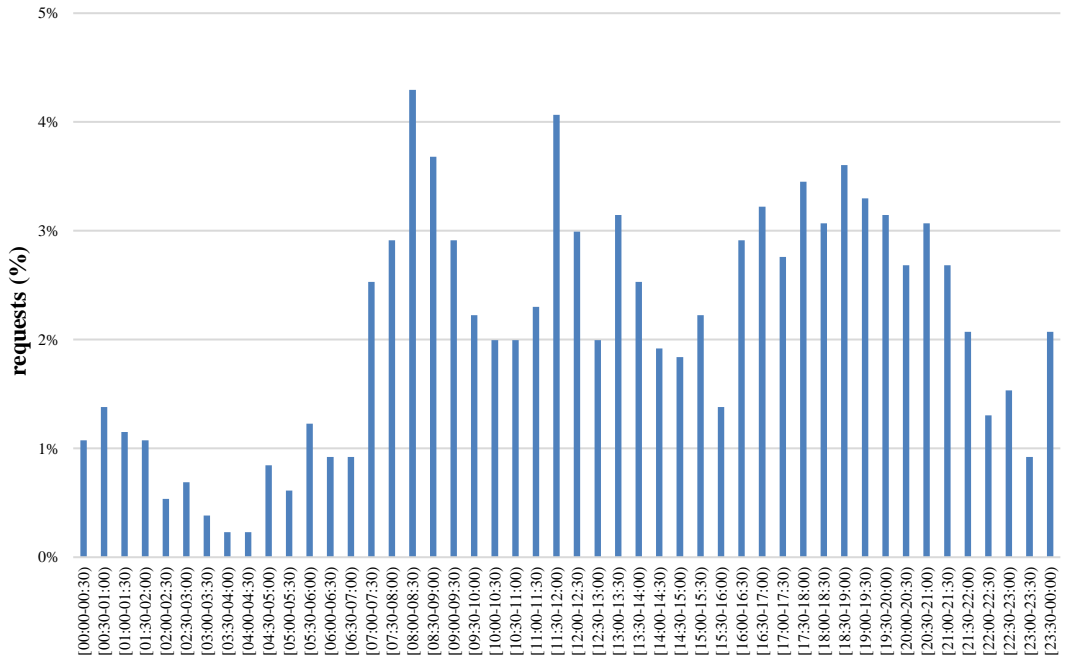


Fig. 8. Requests for half an hour from/to Termini train station

4. Descriptive analysis results

The decision trees technique was chosen to analyse the observed data. Among the different techniques for constructing regression trees, the classification and regression trees (CART) by Breiman et al. (1984) was used. This algorithm is implemented in R software (Quinlan, 1993; Hothorn et al., 2012) and allows data to be portioned into smaller groups that are more homogenous with respect to the response. The technique aim is to identify which predictors are relevant in the description of the phenomenon to obtain reliable models. To achieve outcome homogeneity, regression tree determine the predictor (attribute) to split on and value of the split, the complexity of the tree, the prediction model in the terminal nodes (Khun and Johnson, 2013). For regression, the model starts considering the whole data set and searches each distinct value of each predictor to find the predictor and splits value that partitions the data into two groups such that the overall sums of squares error are minimized.

The dataset used to build the regression tree consists of the total requests for working day from/to the

zone of Termini railway station (more than 1200 trips). Therefore, the dependent variable is the number of taxi requests in the working days in each traffic zone. The predictors (independent variables) considered are the following:

- Population (POP),
- Distance (km), i.e. origin - destination travel distance of journey,
- Travel time (hh), i.e. origin - destination travel time of journey,
- Number of metro stations within the zone,
- PGTU_1 (city centre), dummy variable equal to 1 for central zone, 0 otherwise;
- Metro, dummy variable equal to 1 for central zone, 0 otherwise;
- Number of employees grouped for type of activity:
- Agriculture (AGR),
- Industry and constructions (IND),
- Services (SERV),
- Public Administration (PubAm),
- Instruction, health, artistic activity (IstrSan),

- Other activities (Altro),
- Commerce (wholesale and retail: Com)
- HoReCa, i.e. Hotel Restaurant and Catering,
- Real estate activities (LL),
- Scientific and technical activities (MM),
- Rental, travel agencies, business support services (NN),
- Time slice, dummy variables for the six identified time slices.

The choice of previous predictors is made to consider the possible variables influencing the use of the taxi service according to land use and socio-economic data. Therefore, the main attributes considered are: zonal population, travel attributes (i.e. travel time and distance), the possibility of using an alternative transport mode (for example the metro service), the number of employees (assuming that the use of the taxi may depend on the type of work). The obtained best tree is plotted in fig. 9. From the analysis emerges that the main predictor is related with the professional, scientific and technical activities (MM): the employees in this sector are hence split at the value of 5288. Next partition is in relation to the time slice, if it belongs to the time slice 6 ($FA6 > 0$) or not. Subsequent nodes relate to HoReCa employees: the split value is 1006 for a branch, 703 for the other. In the first case, the construction of the tree stops. In the second case, the considered predictors are the commerce employees (split value equal to 703) and the time slice (if in time slice 5 or not). Next nodes relate with time slice and population. These results confirm the findings obtained in other cities. For example (as in Yang et al., 2018) activities related to professional ones are one of the main producers of taxi requests. On the other hand, time becomes the relevant discriminant: due to the limited transit services in the city of Rome (e.g. the subway stops service at 11.30 pm), in the evening, the set of zones with higher values of requests corresponds to zones with higher number of employees in HoReCa. Such zones produce the higher number of requests during the midday and first hours of the afternoon, i.e. when guests leave or arrive. In the afternoon, the highest values of taxis requests are in zone with higher number of resident and employees in warehouses and retail activities (Com).

Following such indications, a set of regression models have been estimated (eqs. from 1 to 5).

$$TR^{1-2} = -0.000061 \cdot Com + 0.00073 \cdot MM + 0.134 \quad (1)$$

(-3.87)
(5.84)
(4.89)

$$R^2 = 0.20$$

$$TR^3 = -0.000024 \cdot POP + -0.00062 \cdot MM + 0.860 \quad (2)$$

(-3.497)
(8.74)
(3.82)

$$R^2 = 0.54$$

$$TR^4 = -0.000030 \cdot POP + 0.00049 \cdot MM + 0.877 \quad (3)$$

(-4.97)
(7.89)
(4.47)

$$R^2 = 0.51$$

$$TR^5 = -0.000020 \cdot POP + 0.00055 \cdot MM + 0.749 \quad (4)$$

(-4.34)
(10.10)
(4.34)

$$R^2 = 0.61$$

$$TR^6 = -0.000028 \cdot POP + 0.00035 \cdot HoReCa + \quad (5)$$

(-4.41)
(1.57)

+ 0.00036 \cdot MM + 0.601
(3.32)
(3.72)

$$R^2 = 0.51$$

where TR^x is the hourly taxi request in time slice x and (\cdot) is the t -st value. All the calibrated parameters are statistically significant and correct in signs showing similar results of those presents in literature for passenger mobility (Cascetta, 2009). In particular, it can be noted the common and shared contribution of employees on professional, scientific and technical activities (MM). Population (POP) becomes significant for daily trip forecasting (from 07:00 to midnight), while the employees at retail (Com) or touristic activities (HoReCa) are significant from afternoon to first hours of the day.

5. Conclusions

The paper proposed a methodology for investigating and forecasting on-demand service requests exploiting automated vehicle data. Then, the results of demand analysis and modelling related to the taxi service in Rome are presented. The data analysed relies with the taxi trips undertaken in February 2014 (about 756 thousand records, related to 312 taxi drivers on about 7700 taxis in the entire city). In fact, most of taxis operating in Rome are already equipped with mobile devices (GPS), and such a data were available. The taxis are hence monitored remotely, the data collected provide driver id, taxi position, day and time.

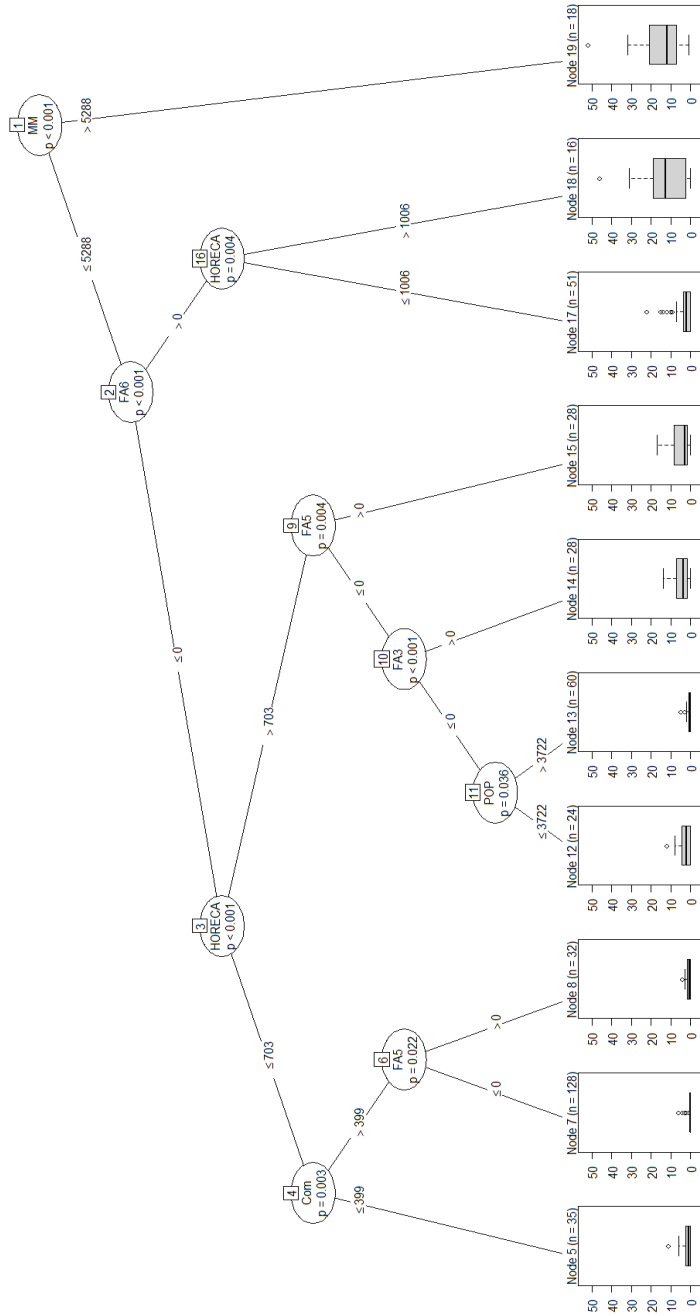


Fig. 9. Basic regression tree

The data analysis implies spatial and temporal considerations. Spatially, it emerged that transport hubs present the large requests: Termini railway station has the biggest attraction for taxi traffic, followed by the Tiburtina station. As far as spatial analysis is concerned, the majority of taxi trips begins and ends within the main road ring (GRA) and from/to Fiumicino's Airport.

Temporally, the distribution for weekday and for half an hour is considered: the analysis suggests a partition of the requests in six time slices. On an average of 230 taxis analysed per day, about 30% of their total trips are taken in the morning from 5am to 8am; while the remaining 70% of their trips are distributed over the rest of the day with two remarkable values at 4pm and at 6pm.

Then, the regression trees technique was used to partition the data and to identify the main attributes to be used in developing regressive models. The dependent variable (response) considers the taxi requests; the predictors (attributes) are a set of representative variables of the considered study area and time slice. It emerged that the root of the tree is represented by the employees on professional, scientific and technical activities. The nodes of the tree are identified in correspondence of the time slices and of HoReCa and Commerce employees. Further analyses are also in progress to improve these first results, including zonal and level-of-service attributes (e.g. travel time and distance), and to develop other models, that can be used to simulate this segment of urban mobility in the ex-ante assessment of new on-demand services.

Acknowledgment

Authors thank Luis Moreira-Matias for suggestions in obtaining data and identifying applied predictive techniques.

References

- [1] AL-AYYASH, Z., ABOU-ZEID, M., & KAYSI, I., 2016. Modeling the demand for a shared-ride taxi service: An application to an organization-based context. *Transport Policy*, 48, 169-182.
- [2] ALONSO, B., BARREDA, R., DELL'OLIO, L., & IBEAS, A., 2018. Modelling user perception of taxi service quality. *Transport Policy*, 63, 157-164.
- [3] ALONSO, B., IBEAS, A., MUSOLINO, G., RINDONE, C., VITETTA, A., 2019. Effects of traffic control regulation on Network Macroscopic Fundamental Diagram: A statistical analysis of real data, *Transportation Research Part A: Policy and Practice*, 126, 136-151.
- [4] BARANN, B., BEVERUNGEN, D., & MÜLLER, O., 2017. An open-data approach for quantifying the potential of taxi ridesharing. In: *Decision Support Systems* 99, 86-95.
- [5] BISCHOFF, J., MACIEJEWSKI, M., & SOHR, A., 2015. Analysis of Berlin's taxi services by exploring GPS traces". In: *Proc. Int'l Conf. Models and Technologies for Intelligent Transportation System*.
- [6] BRACCIALE, L., BONOLA, M., LORETI, P., BIANCHI G., AMICI, R., & RABUFFI, A., 2014. CRAWDAD dataset roma/taxi, downloaded from <https://crawdad.org/roma/taxi/20140717>.
- [7] BREIMAN, L., FRIEDMAN, J.H., & OLSHEN, R.A., & STONE, C.J., 1984. *Classification Regression Trees*, Wadsworth International Group, Belmont, California.
- [8] CAI H., ZHAN X., ZHU J., JIA, X., CHIU, A. S. F., & XUG, M., 2016. Understanding taxi travel patterns. *Physica A: Statistical Mechanics and its Applications*, 457, 90-597.
- [9] CASCETTA, E., 2009. *Transportation Systems Analysis: Models and Applications*. Springer.
- [10] COMI, A., PERSIA, L., POLIMENI, A., CAMPAGNA, A. & MEZZAVILLA, L., 2019. A methodology to design and assess scenarios within SULPS: the case of Bologna. *City Logistics Conference XI*, Dubrovnik, Croatia.
- [11] CROCE, A. I., MUSOLINO, G., RINDONE, C. & VITETTA, A., 2019. Transport System Models and Big Data: Zoning and Graph Building with Traditional Surveys, FCD and GIS. *International Journal of Geo-Information*, 8(4).
- [12] CZECH, P., TUROŃ, K., & BARCIK, J., 2018. Autonomous vehicles: basic issues. *Scientific Journal of Silesian University of Technology*, 100, 15-22.
- [13] FONT, A., GUISEPPIN, L., BLANGIARDO, M., GHERSI, V. & FULLER, G. W., 2019. A tale of two cities: is air pollution improving in Paris and London?, *Environmental Pollution*, 249, 1-12.

- [14] GRAU, J. M. S., & ESTRADA M., 2019. Social optimal shiftsnuzz and fares for the Barcelona taxi sector. *Transport Policy*, 76, 111-122.
- [15] HADAVI, S., VERLINDE, S., VERBEKE, W., MACHARIS, C. & GUNS, T. 2018. Monitoring Urban-Freight Transport Based on GPS Trajectories of Heavy-Goods Vehicles, *IEEE Transactions on Intelligent Transportation Systems*.
- [16] HE, F., & SHEN Z. J. M., 2015. Modeling taxi services with smartphone-based e-hailing applications. *Transportation Research Part C: Emerging Technologies*, 58, 93-106.
- [17] HOTHORN, T., HORNIK, K., & ZEILEIS, A., 2012. Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, 15(3), 651-674.
- [18] KHUN, M., & JOHNSON, K., 2013. *Applied Predictive Modeling*. Springer New York Heidelberg Dordrecht London.
- [19] LI, B., CAI, Z., JIANG, L., SU, S. & HUANG, X., 2019. Exploring urban taxi ridership and local associated factors using GPS data and geographically weighted regression, *Cities*, 87, 68-86.
- [20] LIU, X., GONG, L., GONG, Y., LIU, Y., 2015. Revealing travel patterns and city structure with taxi trip data. *Journal of Transport Geography*, 43, 78-90.
- [21] MACIEJEWSKI M., 2014. Benchmarking minimum passenger waiting time in online taxi dispatching with exact offline optimization methods. *Archives of Transport*, 30 (2), 67-75.
- [22] MOREIRA-MATIAS, L., GAMA, J., FERREIRA, M., MENDES-MOREIRA, J. & DAMAS, L., 2016. Time-evolving O-D matrix estimation using high-speed GPS data streams, *Expert Systems with Applications*, 44, 275-288.
- [23] MOREIRA-MATIAS, L., GAMA, J., FERREIRA, M., MENDES-MOREIRA, J. & DAMAS, L., 2013a. Predicting Taxi-Passenger Demand Using Streaming Data, *IEEE Transactions on Intelligent Transportation Systems*, 14(3), 1393-1402.
- [24] MOREIRA-MATIAS, L., GAMA, J., FERREIRA, M., MENDES-MOREIRA, J. & DAMAS, L., 2013b. On predicting the taxi-passenger demand: a real-time approach. In: Correia L., Reis L.P., Cascalho J. (eds) *Progress in Artificial Intelligence*. EPIA 2013. Lecture Notes in Computer Science, vol 8154. Springer, Berlin, Heidelberg.
- [25] NUZZOLO, A. & COMI, A., 2017. Real-time modelling of normative travel strategies on unreliable dynamic transit networks: a framework analysis, *Modelling Intelligent Multi-Modal Transit Systems*, Nuzzolo, A. and Lam, W. H. K. (eds), CRC Press, Taylor & Francis Group, Boca Raton (FL, USA), 130 – 151.
- [26] NUZZOLO A., COMI, A., PAPA, E., & POLIMENI, A., 2019. Understanding taxi travel demand patterns through GPS data. E. G. Nathanail and I. D. Karakikes (Eds.): CSUM 2018, AISC 879, 445-452, Springer Nature Switzerland AG.
- [27] PIRAS, G., PINI, F. & GARCIA, D. A., 2019. Correlations of PM10 concentrations in urban areas with vehicle fleet development, rain precipitation and diesel fuel sales, *Atmospheric Pollution Research*, 10(4), 1165-1179.
- [28] POLIMENI, A. & VITETTA, A., 2014. Vehicle routing in urban areas: an optimal approach with cost function calibration. *Transportmetrica B: transport dynamics*, 2(1), 1-19.
- [29] QUINLAN, J.R., 1993. *C4.5: Programs for machine learning*. San Francisco, CA: Morgan Kaufmann Publishers Inc.
- [30] RAMEZANI, M., & NOURINEJAD, M., 2018. Dynamic modeling and control of taxi services in large-scale urban networks: A macroscopic approach”, *Transportation Research Part C: Emerging Technologies*, 98, 203-219.
- [31] ROMA CAPITALE, DIPARTIMENTO MOBILITÀ E TRASPORTI (2015). Piano generale del traffico urbano di Roma Capitale.
- [32] RUSSO, F. & COMI, A. (2016). Urban Freight Transport Planning towards Green Goals: Synthetic Environmental Evidence from Tested Results, *Sustainability*, 8 (4), 381.
- [33] SUMAN, H.K. & BOLIA, N.B., 2019. A Review of Service Assessment Attributes and Improvement Strategies for Public Transport, *Transportation in Developing Economies*, 5 (1).
- [34] SUMP (2013). *Guidelines. Developing and Implementing a Sustainable Urban Mobility Plan*; European Commission: Brussels, Belgium.
- [35] TANG, J., LIU, F., WANG, Y., & WANG, H., 2015. Uncovering urban human mobility from large scale taxi GPS data. *Physica A: Statistical Mechanics and its Applications*, 438, 140-153.

- [36] WANG, W., PAN, L., YUAN, N., ZHANG, S., & LIU, D., 2015. A comparative analysis of intra-city human mobility by taxi. *Physica A: Statistical Mechanics and its Applications*, 420, 134-147.
- [37] WONG, K. I., WONG, S.C., & YANG, H., 2001. Modeling urban taxi services in congested road networks with elastic demand. *Transportation Research Part B: Methodological*, 35(9), 819-842.
- [38] WONG, K.I., WONG, S.C., YANG, H., & WU, J.H., 2008. Modeling urban taxi services with multiple user classes and vehicle modes. *Transportation Research Part B: Methodological*, 42(10), 985-1007.
- [39] XIAO Q. & HE R., 2017. Carpooling scheme selection for taxi carpooling passengers: a multi-objective model and optimisation algorithm. *The Archives of Transport*, 42 (2), 67-75.
- [40] YANG, Z., FRANZ, M.L., ZHU, S., MAHMOUDI, J., NASRI, A. & ZHANG L., 2018. Analysis of Washington, DC taxi demand using GPS and land-use data. *Journal of Transport Geography*, 66, 35-44.
- [41] ZHENG, Z., RASOULI, S., & TIMMERMANS, H., 2018. Modeling taxi driver anticipatory behavior. *Computers, Environment and Urban Systems*, 69, 133-141