Optimizing Locations for a Vehicle Sharing System

V. Prem Kumar

Michel Bierlaire

Transportation and Mobility Laboratory (TRANSP-OR) School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), SWITZERLAND-1015 Email: [prem.viswanathan, michel.bierlaire]@epfl.ch

Car-sharing business provides members access to a fleet of shared-use vehicles in a network of locations on a short-term, as-needed basis. It allows individuals to gain the benefits of private vehicle use without the costs and responsibilities of ownership. The primary advantage of car sharing as opposed to the car rental business is that it provides flexibility of using the vehicles for shorter periods of time. Secondly if you are a registered customer of the company, the process of accessing and using the vehicle is quite simple, self-serving and hassle-free.

One of the main problems faced by the car-sharing businesses is that of finding the "best" stations to place the facility. These locations should be chosen by attractiveness of various parameters such as the socio-demographic-economic profile of the population that resides or works at the location and accessibility of other forms of public transport in that location. However the presence of one or more stations in the vicinity and the type of vehicles placed in a station are also critical for decision-making. It must be noted that the attractiveness factor of a parameter would be shared between different stations if they are present in the sphere of influence of that parameter.

This study analyzes the performance of an electric car sharing service across different stations in and around the city of Nice. This service is operated by a major public transport operator in France and commenced its operations in April 2011. The main measure of performance of the stations is the average number of rides (usages) per day. The objective of this study is two-fold – one, to analyze the performance of the car-sharing service across all stations and estimate the key drivers of demand, and secondly, to use these drivers to identify future station locations, such that the overall system performance is maximized.

The methodology used by us to optimize the station locations follows two steps. In the first step, we perform an extensive data analysis and determine all those factors that we expect to be driving the demand for the service and build a linear regression model with station performance as the dependent variable that is explained by a host of other independent variables, such as public transport ridership, share of carusers in the locality, share of high income / education groups in the locality population, population density and presence of other mobility generators such as hotels, colleges and commercial centers.

In the second and final step, we use the attractiveness of the different localities to optimally locate the new stations for the service. The main trade-off decision made by the model involves locating more stations in "highly attractive" localities versus locating new stations in "less attractive" but untapped localities. We validated and reported our results on a set of new stations that were opened quite recently, after the commencement of our study.

1. Introduction

Last century witnessed a major innovation in personal mobility and transportation. While the invention of steam-engine revolutionized public transport, it was the motor car invention in 20^{th} century that wielded the biggest impact on personal transportation. After WW-II, increased incomes and lower production costs resulted in deep penetration of personal cars in middle-class households of the developed countries. However by 1970s and 1980s, the negative impact of heavy usage of personal vehicles to the society and environment became visible. In spite of huge investments in building and maintaining a large road infrastructure, it was realized that the time spent by the society on traffic congestions reduced the availability of quality time for work and leisure. Further, the impact on environment due to CO_2 emissions by the vehicles was observed to be irreparable. While individual mobility needs could not be sacrificed as human society has attuned itself to these needs, one of the ways to curtail the adverse impact on society and environment is by reducing our dependence on personal vehicles. In the recent time, this has generally been achieved by innovative business concepts such as car-pooling and car-sharing. Car sharing concept has a lot of social and individual advantages.

A major benefit of car-sharing at individual level is that it saves money. The cost of vehicle ownership, maintenance, parking and insurance is shared among several users and thus making the service more affordable for every customer. While it may seem counter-intuitive, it is a truth that car-sharing service takes more cars off the roads. Car-sharing may also be helpful to the society in other ways. Some persons may buy bigger vehicles for running occasional trips or errands. But with a membership to a car-sharing service, flexibility of vehicle choice is available to the user depending on the need.

The basic premise of a car-sharing business is that the number of vehicles required to meet the vehicle requirements of several individuals and families is less when they share a single vehicle than when each of them has their own vehicle. It is also believed that individuals use their automobiles only a small portion of each day, as little as an hour or less, during a normal day. According to Katzev (2003), relatively high fixed costs of owning a car are usually ignored when individuals decide whether or not to drive by car. If they think about the cost of travel behavior at all, individuals tend to focus on the low variable costs associated with each trip. This leads them to travel by car more often than they would if they had to pay for each trip, as they do when they drive a car-share vehicle. In this way, once individuals become more mindful of the variable costs of each trip, car sharing is expected to reduce the overall level of vehicle miles of travel. Since members are also required to reserve a vehicle in advance and to a certain extent plan their travel route, they should be less likely to take spur-of-the-moment trips than they are in a privately owned car. Finally, when a private vehicle is no longer available, it is anticipated that car-share members will be motivated to rely more on alternative travel modes, such as carpooling, biking, and public transit, thereby reducing the environmental and social problems associated with automobiles.

European countries have been the early adopters of car sharing systems, with oldest operational systems running since 1980s in Switzerland and Germany. Several published works corroborate the fact that those who belong to the European car-sharing organizations drive considerably less than they did before they had become members. There are primarily two classes of users of the car sharing service. One class consists of a group of individuals who share a fleet of vehicles located within a reasonable walking distance to their residence or workplace. Their usage of cars is primarily for short, convenience-based trips and visits. The second group of subscribers share several cars located at central mobility interchanges such as transits hubs, airports, or rail stations. This class primarily consists of visitors to the city or persons that use shared vehicles for business.

Veolia Transport recently started the Auto Bleue car sharing system around the city of Nice. One of the defining aspects of Auto Bleue service is that the entire fleet of vehicles is electric cars and thus the service has its backing from the local government. Successful operation of this service is hoped to bring down the use of personal cars on the congested roads of Nice. In addition to traffic congestion, parking of personal vehicles is also a major concern in the city. In order to use the vehicles in the fleet, Auto Bleue customers simply telephone the reservation system or book it online. To pick up the car, they can either walk to the nearest station or drive to the station of choice. There are dedicated slots to park other electric vehicles adjacent to the Auto Bleue vehicles. Customers can also reach the station through public transport. On completion of usage, the customer drives the vehicle back to the same station and releases the car for next user. At this stage, three vehicles of one or two types are available across the different stations. While the Peugeot Ion and Mia models are considered as family or personal vehicles, Berlingo cars are usually available in the fleet to give members an efficient way to meet infrequent needs, such as hauling, moving, and transporting large groups. Auto Bleue bears all of the costs of vehicle maintenance, service and repairs, insurance coverage, parking, and the cost of recharging.



Fig 1: White, Yellow and Pink markers represent the chronological phases of locating stations with white ones in Jan / Feb 2012, yellow ones in Sep 2011 and pink ones in Apr 2011

Even though Auto Bleue system commenced less than a year ago, it has undergone very rapid expansion and also has phenomenal growth plans for the near future. The growth so far has been phased and conducted in three phases, viz., Phase 1 (in April 2011), Phase 2.1 (Sep 2011) and Phase 2.2 (Jan-Feb 2012). Since this study commenced before the Phase 2.2 stations were

already commissioned, we have excluded them in our modeling. However the performance of the new stations over the brief period till the end of the study was used to validate the performance of our models. The growth of the stations in the three phases so far is pictorially represented in the map (Fig 1).

Currently there are around 40 stations and three vehicles at each station in the car sharing fleet. However the service is being expanded at a fast pace and it is expected to increase to 68 stations and 204 vehicles in the near future. Thus, one of the primary objectives of this project is to study the choice of and recommend possible locations for locating the next 28 stations and 84 cars so that the overall facility utilization and service reach is maximized.

Given this exceptional growth, one of the major concerns for the car sharing service, at this point, is the selection of ideal locations to buy space and place the facility. Prior research on the car-sharing businesses has revealed that one of the critical factors impacting the service performance is the location of stations. As most of the users tend to walk down to the nearest station to pick up the vehicles, it is important to locate the facility as close to the user as possible. At the same time, it is necessary that each station is a catchment for a substantial number of users so that the profitability of the system is not compromised. If we look at the performance (average number of bookings per day) of the stations commissioned in the first two stages, there is no obvious visible trend or pattern (refer Fig 2).



Fig 2: Stations with blue marker are among the top one-third, green represent the middle onethird and red represent the bottom one-third in terms of performance

By simply looking at the map above, it is difficult to give a reason for the performance of the stations by considering their geography. In general, the stations within the city perform better

than the sub-urbs, but low performing stations within the city, or high performing sub-urbs, cannot be explained. The selection of locations so far has not followed any scientific approach but relied on the gut-feel of the decision-maker. One of the major challenges ahead for Auto Bleue business will be to manage the upcoming investments prudently with an analytical and scientifically driven decision making process to locate the next set of stations.

Thus we summarize the key objectives of the study as the following:

- Understand and analyze the various drivers of demand for Auto Bleue service from the rich data that is made available to us from the different sources,
- Build a mathematical model to represent the performance of Auto Bleue station through these drivers, and
- Use the mathematical model to optimize the location of future stations.

In the subsequent sections of this report, we would aim to analyze the performance of the different Auto Bleue stations and identify the factors contributing the performance of the stations and use these factors to identify optimal locations for new stations. In the section 2 of this report, we would analyze the research literature available on this topic. In section 3, we will briefly describe the mathematical methodology that we would employ to optimally locate the stations. In section 4, we will discuss the results of the methodology and provide our recommendations. Section 5 will conclude the report with a future outlook and other areas that can potentially looked into to improve the overall utilization of this service.

2. Literature Survey

Shaheen et al. (1998) presented one of the first literatures on car sharing business. This article analyzed and compared the trends of car sharing services in US and Europe and suggests that technology and flexibility in the services is the way forward for car-sharing service organizations. However, an important input provided by the article from the point of view of locating stations is that the car-sharing service organizations are likely to be more successful economically if they provide a dense network and variety of vehicles, serve a diverse mix of users, create joint marketing partnerships, design a flexible but simple pricing system, and provide for easy emergency access to taxis and long-term car rentals.

Ciari et al. (2008) use an agent-based traffic micro simulation tool to show that the benefits of car-sharing service to transportation system, society and environment are dependent on the capillarity of system, its flexibility and its integration with other mobility tools. Efthymiou et al. (2012) use a clustering approach with a number of variables, such as the population characteristics, points of interest and the characteristics of the electric utility grid around the candidate positions, to suggest potential locations for placing the stations. The driving idea behind the approach is to weigh the variables in order to produce a single spatial map to support the decision-makers.

Uesugi et al. (2007) consider the one-way car-sharing system and suggest a simulation based optimal vehicle assignment at the stations so that the distribution balance of parked vehicles is maintained. Correia and Antunes (2012) present an optimization approach to depot location in one-way car-sharing systems where vehicle stock imbalance issues are addressed. Their approach is based on mixed-integer programming models whose objective is to maximize the profits of the car-sharing organization considering all the revenues and costs involved. The practical usefulness of the approach is illustrated with a case study involving the municipality of Lisbon. The results we have obtained from this study provided a clear insight into the impact of depot location and trip selection schemes on the profitability of such systems.

Fan et al. (2008) consider the dynamic vehicle allocation problem, in a car-sharing context, as a decision-making problem for vehicle fleet management in both time and space to maximize profits for the car-sharing service operator. A multistage stochastic linear integer model with recourse is formulated that can account for system uncertainties such as car-sharing demand variation. A stochastic optimization method based on Monte Carlo sampling is proposed to solve the car-sharing dynamic vehicle allocation problem.

The brief survey of literature on the car-sharing systems tends to show that most of the methodological research is published in the last 3-4 years. As with the early stage research on any topic, the growth of research in this fleet is fairly hap-hazard. While a lot of researchers are publishing their results for a highly complex one-way car sharing system, even the optimization of simple return-trip car-sharing system appears to be far from complete. Most of the return-trip car-sharing systems tend to focus on using the drivers of demand to identify clusters to locate stations or consider the uncertainty to demand to optimize the fleet size at the station. In this context, our paper aims to plug the gap between the existing literatures in the following ways:

- 1. When the drivers of demand are generally known, locating stations is a trade-off between the "attractiveness" of the location and the number of existing or proposed stations in the vicinity. It is well-known that presence of too many stations in the same locality is likely to cannibalize the demand at the locality, while a "non-attractive" location to place a station may not meet the needs of the station. This trade-off evaluation has virtually remained unexplored in the literature.
- 2. We present a mixed-integer quadratic (higher-order) assignment formulation and propose a heuristic to solve the same.
- 3. Most of the published works are not applied to the real-life scenario and one of our contributions of this study is to apply the results of our methodology to validate it with real data as obtained for Auto Bleue service in the city of Nice.

3. Brief Data Analysis

As we have already seen, it is not possible to explain the performance of the stations only with their geography. In this section, we would dive deeper into the different data sources and attempt to explain the performance of the stations with these data elements. The different data sources that we will study during the course of this project are as follows:

- Station location data
- Auto Bleue customer data
- Auto Bleue booking data
- Auto Bleue trip data
- Velobleue (Bicycle sharing sister-company) data

- Public transport data
- Mobility attractors data (presence of hotels, shopping complexes, etc. in the vicinity)
- Demographic data from the Nice Cote d'Azure (NCA) provided at IRIS level

3.1 Auto Bleue Data

Let us first look at the Auto Bleue usage data. In particular, we feel that it might be interesting to explore the length of the trips when an Auto Bleue vehicle is being used. This is particularly important to see since Auto Bleue runs a fleet of electric vehicles which needs to be recharged after running about 100 KM. Moreover, car sharing services are designed to optimize the prices for short duration trips, typically up to half-a-day, beyond which normal car rental services become cheaper. Length of usage of Autobleue bookings is shown in Fig 3.



Fig 3: Length of usage of the Auto Bleue vehicles during one booking

Nearly 75% of the bookings are for duration of less than 4 hours and almost 88% are for less than 6 hours. Let us now direct our attention to the Auto Bleue customer data. Fig 4 gives a break-up by age profile of Auto Bleue customers. Almost 70% of the customers are in the 26-55 age group. The actual profile of the population of NCA has slightly higher proportion of younger persons in the age-group of 18-25 as well as 56+ (Fig 5).



Fig 4: Share of Auto Bleue customers by age



Customer data also reveals that the split by gender is skewed heavily towards men. According to literature, it is in line with many other car sharing services. However we feel that Auto Bleue must run some special promotions to attract corporate customers, whose share in the existing customer category is very less. Fig 6 summarizes the customer profile by gender.



Fig 6: Share of Auto Bleue customers by gender

Further analysis of the customer data reveals that 45% of the customers have used the system very occasionally. We classify these users as "test users", who are probably attracted to electric cars and want to try them on a few occasions and then return to their personal cars. Among the rest of the customers, almost one-third can be classified as occasional customers, one-third as regular customers and the remaining one-third as frequent customers. All customers that use the system thrice or more are referred to as "subscribers" as opposed to "test users". It would be interesting to note that 55% of subscribers (as opposed to 45% test users) contribute to over 93% of all vehicle pickups.



Fig 7: Breakup of pickup bookings closest to a customer's residence



Fig 8: Breakup of pickup bookings by distance to customer's residence

Next we determine the percentage of customers that actually pick the vehicle from a station closest to the residential address as reported to us in Fig 7. The only concern here is that the data available to us does not provide us with the workplace address of the customer and there is no guarantee that the residential address. Therefore, in Fig 8, we present the break-up of vehicle pickups by the distance from customer's residence. It is observed that 58% customers pick-up a vehicle closest to their residence, while as much as 88% pick the vehicle from a station within reasonable walking distance from their residence (up to 1.5 KM). It also indicates that a large number of customers are willing to walk to the second or third nearest station to pick up the vehicles when their preferred vehicle is not available at a station.

When the vehicle is not picked up from the station closest to home, we will now evaluate if there is a pattern in the choice of stations. Fig 9 represents the percentage of booking by customers for whom the station is not nearest to their residence. While there is no distinct pattern among the stations with a high number of such bookings, except may be Augusta Gal (presence of a shopping complex) and Gare Thiers (major mobility interchange), there is clearly pattern among stations where the number of such bookings is less. All of these locations are in the suburbs, which makes notional sense.



Fig 9: Percentage of bookings by station made by customers for whom the station is not closest

Next we look at the tendency of the customers (that are subscribers) to pick up the vehicle from the same station in Fig 10. Interestingly, only a quarter of customers pick up the vehicles from the same station. This is particularly true for customers located in sub-urban locations. About 20% customers pick the vehicle from two stations, while almost 55% customers pick up a car from as many as three or more stations.



Fig 10: Tendency of subscribers to pick vehicles from different stations

We now consider the trip data for Auto Bleue vehicles. When a vehicle is being used by customer, it can potentially make two types of stops during its trip – implicit and explicit. An explicit stop refers to a situation when the car is locked by the customer on a permanent basis using the card. All other stops are referred to as implicit stops. We examine the stops undertaken by the vehicles during their trips and identify the stops that are particularly made close to an existing Auto Bleue station (within 500 m) and shown in Fig 11. 74% of the stops are usually made close to an existing Auto Bleue station.



Fig 11: Location of stops during a trip

Further analysis of the vehicle trip data reveals that a large number of trip stops during the vehicle reservation is made near Auguste Gal and Notre Dame stations. These two stations are located close to shopping and pedestrian locations of the city.

3.2 Velobleue Data

Let us now focus our attention to the data received from Velobleue – which is the sister company of Auto Bleue and shares bicycles around the city of Nice and sub-urbs. Velobleue service has been operating the city and the sub-urban areas since July 2009 and it, like Auto Bleue, has undergone several phases of growth. The service can now be assumed to be fairly stabilized in terms of usage, even though we expect seasonal variations in the pickup of bicycles, unlike that of a car. There is a reason to believe that the performance of Velobleue at certain stations could be a marker for the performance of Auto Bleue in the nearby stations.

The methodology employed by us to relate the Velobleue pickup data with Autobleue pickup data is the following. Pickup data for Velobleue is available till second quarter of 2011, while Auto Bleue data is available till Dec 2011. We first determine the average number of pickups at all Auto Bleue and Velobleue stations for the entire duration of their existence. Then, for every Auto Bleue station, we determine the nearest Velobleue station within 500 m, if there exists one. There usually always exists a Velobleue station near an Auto Bleue station in the city of Nice (though not always in sub-urban areas) as the number of Velobleue stations (175) is much higher than Autobleue stations (29). In many cases, there are many Velobleue stations close to an Auto Bleue station and the gap between the nearest station and the second nearest station is marginal. For example, the nearest Velobleue station to an Auto Bleue station could be at 180 m, but the second nearest one may be at 185 m. Sometimes the performance of these two nearby Velobleue stations vary substantially and, thus, it is not completely fair to compare the Auto Bleue performance with respect to only the nearest Velobleue station. To address this problem, we first consider the nearest Velobleue station within 500 m of an Auto Bleue station. Next we consider the second nearest Velobleue station within 400 m and the third nearest station within 300 m. Thus, for every Auto Bleue station, we consider up to three Velobleue stations in the vicinity. We progressively reduce the area of influence for second and third nearest distances to account for the principle of diminishing returns. We now compute the corresponding Velobleue performance for an Auto Bleue station by weighing the pickups at all these (one, two or three) stations inversely to their distance from the Auto Bleue station. Similarly if there are multiple Auto Bleue stations corresponding to one Velobleue station, the pickups of the Velobleue station is shared inversely proportional to its distance from the corresponding Auto

Bleue stations. The computation of Velobleue performance corresponding to every Autobleue station is pictorially represented in Fig 12.



Fig 12: Computation of the impact of Velobleue performance for each Auto Bleue station

The plot of the performance of the two measures indicates that there is some correlation between the two.

3.3 Public Transportation Data

Like Velobleue, the penetration and usage of public transportation in a locality is expected to be a major driver for the performance of Auto Bleue. It makes intuitive sense because the tendency of use public transportation by a community is indicative of their willingness to accept other shared modes of transport – such as Auto Bleue. For the evaluating of the impact of public transportation on Auto Bleue, we conduct a similar analysis as was done for Velobleue. We have access to the data of bus and tram stops along with the number of rides associated with that stop daily. We relate the Auto Bleue performance at a station with the public transport performance corresponding to a close-by bus or tram stop.

Instead of finding the nearest bus or tram stop to an Auto Bleue station, we sort the stops by the number of rides. We share the rides associated with a particular bus or tram stop within 500 m of an Auto Bleue station to the corresponding station. If there are more than one station, the associated public transport rides is shared between all these stations. This process is done for every bus or tram stop in our data and the impact of public transport ridership is accumulated to the corresponding Auto Bleue station. Similar to Velobleue, we would have the tendency for public transport ridership corresponding to each Auto Bleue station, which would be regressed with the Auto Bleue performance for that station. The correlation between public transport ridership and Auto Bleue performance for each station is significantly stronger than the correlation between Auto Bleue and Velobleue. However correlation improves significantly if it is computed separately for the city and the sub-urban regions.

3.4 NCA Data

We now consider the data obtained from the Nice Cote d'Azure Authorities to see if some of them contribute to the performance of Auto Bleue. This data divides the city of Nice into 146 smaller IRIS communes and 69 sub-urban communes. IRIS data provides commune-wise information on a number of parameters such as population breakup by gender and age groups, occupation and transportation needs for the workplace (such as two-wheelers, personal car, public transport, etc.) The IRIS data is provided as GIS files and it is possible to estimate the perimeter, area, centroid and other geometrical parameters associated with each IRIS commune. We assume that the parameters associated with the commune are uniformly applicable to the entire commune as a whole and does not vary within the same commune.

It might be interesting to understand the layout of the city of Nice with respect to its suburbs. This is pictorially represented in the Figs 13 a and b. Not surprisingly, the size of the communes within the city is fairly small, while those in the sub-urbs are fairly bigger in terms of area. The reason is that the IRIS communes are defined in such a way that the number of people residing in a commune is similar.



Fig 13a: Projection of NCA on google maps



Fig 13b: Nice city (blue) versus sub-urbs

Using the available data, we now determine the share of the different parameters associated with an IRIS commune such as:

- Share of the individuals that drive to workplace on their personal cars
- Share of the individuals that use a two-wheeler to reach the workplace
- Share of individuals that use public transport to reach their workplace
- Share of individuals that do not require a transport to reach their workplace
- Share of men in the population
- Share of different age groups in the population
- Population density

- Share of entrepreneurs / craftsperson among residents
- Share of managers / professionals among residents
- Share of employees / associate professionals among residents
- Share of workers among residents

The next step would be to compare these parameters with the Auto Bleue performance at the corresponding station. While it may be tempting to use the Auto Bleue performance of station and compare the same with the value of the parameter associated with the IRIS commune on whom the station is located, but it could be incorrect. The main reason is that the presence of an Auto Bleue station at an IRIS commune (especially within the city where the areas of these communes are quite small) does not only attract the residents of that commune. The method employed to compute the value of parameter corresponding to an Auto Bleue station is represented pictorially in Fig 14.



Fig 14: Calculation of the parameter value corresponding to an Auto Bleue station

In the figure shown above, Autobleue station is located in IRIS 1, but its catchment area spreads into three other IRIS communes. For example, if we want to compute the car ridership to workplace corresponding to the Auto Bleue station shown above, we first draw a circle of influence (shown in green) of the station, 200 m in this case. We then compute the areas covered by this circle under the different IRIS communes. We also know the car ridership parameter for each of the commune separately. To evaluate the car ridership around the Auto Bleue station, we multiply the area covered by the circle with the car ridership of the commune and cumulate over all the "influenced" communes. This value is then divided by the area of the commune to obtain the car ridership parameter for the station. This method is employed to compute the different parameters corresponding to each Auto Bleue station and then correlated with the performance of the station to see if that parameter is a driver for the Auto Bleue demand at that station. For example, results of such a correlation for car ridership to workplace are shown in Fig 15.



Fig 15: Correlation between car ridership to workplace and Auto Bleue performance

As expected, co-relationship between car ridership and Auto Bleue usage is fairly strong on the negative side. Similarly, the correlation between the share of high income, educated groups (managers / professionals) and the performance of Auto Bleue usage is positive. We determine the correlation between other parameters associated with NCA demographics and Auto Bleue performance and check if they could be reasonable drivers for demand.

3.5 Other Mobility Attractors Data

We now look at other generators of mobility and see if their presence around an Auto Bleue station drives the demand for vehicles at that station. The typical mobility attractors considered by us during the course of our study are Hospitals, Colleges, Cultural Centers, Shopping complexes, etc. We consider the impact of any mobility attractor institution on the Auto Bleue performance at the station if its location is within 500 m of the station and closest to it. Figures 16, 17, 18 and 19 represent the impact of the mobility attractor institution of a nearby Auto Bleue station. The measure of impact is the number of footfalls corresponding to the mobility attractor.





Fig 17: Hotels near Auto Bleue stations



Fig 18: Cultural Centers near Auto Bleue



Fig 19: Hospitals near Auto Bleue stations

This concludes our study of the different parameters associated that could potentially influence the performance of Auto Bleue. In the next section, we will build integrated mathematical models that relate these drivers to the Auto Bleue performance and also make suggestions to select the new locations for the future stations such that the overall system performance and usage is optimized.

4. Mathematical Model

We will now aim to build a mathematical model that would first identify the drivers of the Auto Bleue performance and then use these drivers to optimize the selection of future locations. We plan to achieve this through a two-step process, wherein we first build a multi-linear regression model with Auto Bleue performance as the dependent variable and then use the identified drivers to optimize the location in the second step.

4.1 Linear Regression Model

Regression analysis is a technique for modeling and analyzing several variables, when the focus is on the relationship between a dependent variable and one or more independent variables. In our case, the dependent variable refers to the Auto Bleue performance (average daily pickup) at a particular station and the independent variables are the set of drivers which showed reasonable correlation with the Auto Bleue performance. The one of the main assumption related to our regression model is that the relationship between the Auto Bleue performance and these independent variables is linear in nature. While assumption of linearity is fairly strong and restrictive, we still pursue it for its ease and simplicity in measuring performance.

We will estimate and select the following independent variables corresponding to each Auto Bleue station for our linear regression model. Eventually we will skip some variables if the model fit is poor.

- Velobleue performance indicator
- Public transport ridership
- Share of residents using their personal cars for transport to office
- Share of residents using two-wheeler or public transport to reach workplace
- Share of residents that are entrepreuners / craftsmen

- Share of residents that are Managers / Professionals
- Share of residents that are employees and associate professionals
- Share of residents that are workers
- Population density
- Share of males in the population
- Share of 25-54 age group persons in the locality
- Special variable for Gare Thiers
- Distance to the nearest Auto Bleue station
- Commercial center
- College Lycee
- Hotels
- Hospitals

In sections 3.2 to 3.5, we have explained how the values for certain parameters from Velobleue, Public Transport, NCA and mobility attractors data can be associated with an Auto Bleue station. The only difference in the computation relates to the circle of influence of Velobleue and Public Transport. We consider the impact of a Public Transport stop on all Auto Bleue stations within a radius of 500 m from it. Also for simplicity, the impact of public transport ridership gets equally divided among all these stations as shown in Fig 20.



Fig 20: Circle of influence of public transport ridership on all Auto Bleue stations

We also introduce a special binary variable for Gare Thiers to account for two features. One, Gare Thiers was the first station to be inaugurated by Auto Bleue and it was the only station holding a billboard and thus advertising the service. Secondly, the only agency where customers can walk-in to apply for an Auto Bleue membership is also located at Gare Thiers. For these reasons, we believe that the performance at Gare Thiers can be expected to be higher than normal. Distance between stations is a very important parameter in Auto Bleue business. It is a well-documented that as the inter-distance between the stations get smaller, the catchment area of the stations get impacted and there could be potential cannibalization between the stations. From a business point of view, it means that Auto Bleue keeps adding new stations in the same locality without adding new customers. This can also be gauged from the analysis in section 3.1 where we see that a large portion of Auto Bleue customers have a tendency to book cars from multiple locations, quite often more than three locations. Most studies show that the inter-station distances do not have a linear impact on the performance of the stations. We continue with our theory of circle of influence to measure the impact of distance. Thus, for every Auto Bleue station, we determine the number of other Auto Bleue stations within 500 m and use it as an input for the distance.

We build two separate models for entire booking data and only subscriber booking data. The idea behind building two models is that we want to determine separately the drivers that impact the regular users of the system. We build separate models for the sub-urban parts of the city as the drivers for demand in these parts can be significantly different owing to differences in demographics and transportation needs patterns.

The drivers that impact the Auto Bleue performance are identified as the following:

- Weighted share of high income, high education population (+)
- Weighted share of public transport ridership (+)
- Weighted share to car usage to reach workplace (-)
- Presence of commercial center (+)
- Presence of hotels (+)
- Presence of College Lycee (+)
- Population density (+)
- Distance (-)
- Special Gare Thiers variable (+)

Signs of selected variables in the model are correct. The two models are fairly robust and have reasonable measures of fitness. In addition, the significance of most of the variables selected by the model has good levels of confidence. Presence of collinearity between independent variables has been checked. Standard deviation of error is reasonable.

Let us now drive our attention to the model for the suburban areas. Given that there are so few locations with an Autobleue station in the sub-urban areas, we would not have the degrees of freedom to fit the model with many variables. Since Lindbergh station is geographically close to sub-urban stations, we also include it in our model. The results of the sub-urban model indicate that the only relevant variables to represent the Auto Bleue performance are Weighted Public Transport and Weighted Population Density.

A detailed map representation of some of the variables such as presence of high income / education groups, car ridership to workplace and public transport ridership are shown in the Appendix. This concludes our exercise of building the linear regression based demand drivers model to capture the Auto Bleue performance. In the next section, we would use these drivers to build an optimization model that ensures a reasonable choice of future stations.

4.2 Optimization Model

The purpose of the optimization model is to ensure that the trade-off between the attractiveness of drivers in a locality and its proximity to an existing station is balanced. Having too many stations at attractive locations does not increase the overall system performance. Optimization algorithm is expected to aid the decision-making process by making us look at less attractive locations, which we might have ignored otherwise. Note that the result of the regression model obtained in section 3.1 can be written in the following manner:

$$\begin{aligned} \mathbf{Y}_{\text{locality}} &= \beta_{\text{INTERCEPT}} + \beta_{\text{ME}} \, X_{\text{ME,locality}} + \beta_{\text{PT}} \, X_{\text{PT,locality}} + \beta_{\text{CR}} \, X_{\text{CR,locality}} + \beta_{\text{CC}} \, X_{\text{CC,locality}} + \beta_{\text{Hot}} \\ & X_{\text{Hot,locality}} + \beta_{\text{CL}} \, X_{\text{CL,locality}} + \beta_{\text{PD}} \, X_{\text{PD,locality}} + \beta_{\text{Dist}} \, X_{\text{Dist,locality}} + \beta_{\text{GT}} \, X_{\text{GT}} + \epsilon \end{aligned}$$

where

Y_{locality} represents the performance of an Auto Bleue station at that locality, $\beta_{\text{INTERCEPT}}$ is the intercept obtained from the regression model, β_{ME} is the coefficient of the variable corresponding to Managers and Experts, X_{ME.locality} is the share of Managers and Experts in the population for the locality, β_{PT} is the coefficient of the variable corresponding to Public Transportation ridership, $X_{PT,locality}$ is the value corresponding to the public transport ridership for the locality, β_{CR} is the coefficient of the variable corresponding to personal car ridership to workplace, X_{CR.locality} is the share of personal car ridership to workplace for the locality, β_{CC} is the coefficient of the variable corresponding to presence of commercial center, X_{CC.locality} is the footfalls from commercial center for the locality, β_{Hot} is the coefficient of the variable corresponding to hotels and temporary accommodation, X_{Hot,locality} is the number of rooms in hotels and temporary accommodation for the locality, β_{CL} is the coefficient of the variable corresponding to presence of college lycee, X_{CL,locality} is the number of students in college lycee for the locality, β_{Dist} is the coefficient of the variable corresponding to presence of other Auto Bleue stations, X_{Dist.locality} is the number of other Auto Bleue stations in the locality within 500 m, B_{GT} is the coefficient of the variable corresponding to special Gare Thiers variable, X_{GT.locality} is 1 if the locality is Gare Thiers, and 0 otherwise, ϵ is the error term in the computation of Y

By definition of the linear regression model, error terms have a mean of zero. Thus the expected performance of an Auto Bleue station for a locality depends entire on the X-variables. We are given a set of IRIS communes and their demographic profile in the NCA data. We would exploit this information to locate the next set of stations. By this idea, every locality can be considered as an IRIS commune and the centroid of the locality is a potential location to place a station. While it is not necessary that the stations are actually placed on the centroids of the IRIS communes, but we use the centroid to make measurements of the parameters, such as public transport and distance. Obviously there is an underlying assumption that centroid is representative of every point within the IRIS commune, which may be incorrect, especially for locations on the fringes of the centroid. But we need some assumption to mathematically represent the potential stations and we feel that this would be the most sensible one.

Before we explain the mathematical model, please note that most of the variables that represent the performance of Auto Bleue station are dependent **only** on the locality under consideration. Examples of such variables are share of Managers and Experts in the locality, share of personal car ridership to workplace in the locality, footfalls at the commercial centers in the locality, number of rooms in hotels and temporary accommodation in the locality and the number of students in college lycee for the locality. However, some other variables such as Public Transport ridership and Distance variables depend on other localities as well. Remember that public transport ridership value is measure by dividing the number of rides from all stops within 500 m with the number of Auto Bleue stations in its circle of influence. Thus, this variable would also depend on the presence of other Auto Bleue stations in the nearby localities. Distance variable is computed based on the number of stations in the circle of a given locality. Thus we formulate our optimization model in the following manner:

Let the set of potential locations be k such that $k \in K$

Let k' represent a subset of locations in the vicinity of k (within 500 m) such that $k' \subset K$, $k \notin k'$ Let $s \in S$ be the set of all public transport stops and r_s be the number of rides from each of them Let s_k be the set of stops in the circle of influence of potential station, k, such that $s_k \subset S$ Let k_s be the set of localities in the circle of influence of stop s, such that $k_s \subset K$ Let z_k be a binary variable such that $z_k = 1$ if locality k is selected as a station and 0, otherwise Let the number of localities that need to be selected be n where $n \leq |K|$ Let the expected performance of Auto Bleue at a locality be represented as $Exp(Y_k)$, where $Exp(Y_k) = \beta_{INTERCEPT} + \beta_{ME} X_{ME,k} + \beta_{PT} X_{PT,k} + \beta_{CR} X_{CR,k} + \beta_{CC} X_{CC,k} + \beta_{Hot} X_{Hot,k} + \beta_{CL} X_{CL,k} + \beta_{PD} X_{PD,k} + \beta_{Dist} X_{Dist,k}$

The main objective of our mathematical model is to maximize the combined performance of all the stations, which can be written as:

$$Max \sum_{k=1}^{K} z_k Exp(Y_k) \qquad \dots (1)$$

subject to:

$$\sum_{k=1}^{K} z_k = n \qquad \dots (2)$$

$$X_{Dist,k} = \sum_{k'} z_{k'} \quad \forall k \qquad \dots (3)$$

$$X_{PT,k} = \sum_{s \in s_k} \left(\frac{r_s}{\sum_{k \in k_s} z_k} \right) \qquad \forall \ k \qquad \dots (4)$$

The mathematical formulation (1) - (4), which is an integer non-linear program, is difficult to be solved with a normal solver. However there are several parameters within this model that are fairly independent of the complex portion, which is (4). We therefore use a heuristic to solve this problem, which first estimate the "best" set of Ys based on all parameters except distance and public transport ridership. Remember that this portion of the computation is a constant and is independent of number of operational stations. In the first iteration, we assume that all stations would be operational, i.e., $z_k = 1$, for all k, and compute the contribution of public

transportation and distance. In the next step, we pick the n best locations to place the stations, based on the objective function. Now we recomputed the public transportation and distance contribution assuming that only these n proposed and original stations are operational. Based on the changes in the objective function, we again pick the n best locations to place the stations. This process is repeated iteratively until the selected set of n stations does not change.

We will now validate our models and discuss the results produced by our algorithms in the subsequent section.



Fig 21: The darker shades of blue represent the best localities and unshaded ones are the worst

4.3 Model Validation

In the two sections above, we have presented several maps highlighting the "attractiveness" of the different localities within and outside the city of Nice. The measure of attractiveness is based on mathematics and it is important to validate our approach and models. It is indeed fortunate for us that Auto Bleue underwent a phase of expansion during the course of this study at the last week of January. These stations could not be considered in our model because of their late commissioning and absence of bookings data. Moreover, these stations are expected to achieve stability in performance over the course of next few months.

However the performance of the stations, commissioned in January, till the end of March can be considered as useful inputs to validating our mathematical models. Even though the performance of these stations has not achieved stability, their relative performance with respect to each other would be a favorable indicator for comparison. Of the 13 stations planned to be opened in this phase, 11 are already operational and we have access to the booking data at these stations since their date of opening. We will now compare the performance of these 11 stations with respect to the prediction of the model:

- Auber: This station is located just 50 m from Gare Thiers. It is expected to perform very well, though it is also likely to cannibalize some demand from Gare Thiers as per the model. On the ground, this station performs the best among the 11. Thus the model prediction and performance of the station match.
- Vernier: This station is located north west of Gare Thiers in the light blue region. It is expected to perform well as it is also located at a reasonable distance from other stations. On the ground, this station performs the second best and, thus, the model prediction and performance of the station match.
- Rene Cassin: This station is located north east of Lindberg and close to the airport. It is located in the non-shaded region and hence expected to perform very bad. However, contrary to the prediction of the model, this station is third best and a surprise. Lindberg has certainly experienced some amount of cannibalization since its opening, but it is not difficult to explain the higher performance of this station. Incidentally, Veolia's main office is adjacent to this station and it is a conscious decision by the head of this office to use Auto Bleue vehicles for their official purposes.
- Stalingrad: This station is located along the port of Nice. It is located in the middle of the white-shaded region, but reasonable far from all other existing stations and hence expected to perform marginally. This station is currently the fourth-best and its slightly better performance could be attributable to the specific feature of this commune as the population density would have been underestimated in our model. It must be noted that about 50% of the commune is part of the sea, but the computation of area and population density using the GIS software could not consider this aspect.
- Chateauneuf and Grosso: These two stations are located north-west and south-west of the Saint Philippe station and are currently fifth and sixth best in terms of performance. While both these stations are located in green shaded regions, Chateauneuf is exposed to a light-blue shaded commune in its catchment area. However Grosso is reasonably far from most of the existing stations. Thus the performance of the two stations is very much in line with their expected performance through the mathematical model.
- Napoleon III, La Bornala, Colomars and Grenoble: These stations are all located in the western fringes of the city. These regions are all unshaded and hence expected to perform poorly. These stations also perform poorly on the ground so far and hence in line with the model expectation.
- Georges Ville: This station is located between Philibert and Augusta Gal and west of Blanquai. This station is located in the high performance region, at the junction of dark blue, light blue and green shaded localities. However, three other stations are located within 500 m of this station, even though the distance to the nearest station is more than 300 m. Thus the expected performance of the station as per the model is high, but the station is currently among the bottom five stations. Therefore, it can be said that the model prediction for this station does not match with its actual performance. This is probably explained by the fact that the model is under-estimating the impact of distance between stations and presence of nearby stations.

Fig 22 compares the results of the model with their actual performance for newly commissioned Auto Bleue stations. The figure is divided into four quadrants, such that the straight line along bottom-left corner to top-right corner represents the situation when model prediction is exactly same as the actual performance. Results show that eight of the eleven stations are along this straight line, indicating that model prediction is same as actual performance. For Stalingrad, model predicts slightly lower performance than actual, while the results for Rene Cassin and Georges Ville are significantly different. Overall, our model predicts the performance of new stations reasonably well.



Actual Performance -

Fig 22: Model Validation Results

It must be noted that the "attractive" locations are almost entire based within the city of Nice or in the south sub-urbs of the metropolitan. Northern parts of the Nice metropolitan is, thus, not well-suited and would not be discussed in detail for the choice of locating new stations. This concludes our section of mathematical modeling and we will now summarize the key recommendations of this report in section 5.

5. Model Results

This section will study the different localities within and outside the city of Nice and make recommendations to position new Auto Bleue stations so that the service utilization is maximized. By far, we will go by the suggestions of the mathematical model, but we will also apply intuition where we do not have conclusive evidence to prove the impact mathematically.

Auto Bleue performance around the city center has generally been better than the performance around the fringes of the city center. Some stations, such as Deloye and Ponchettes, did not observe the kind of performance that is expected from those locations. Some other stations, such as Californie and Cyrille Besset, performed much better than expected. While it is easy to capture and relate the performance of the stations on the different parameters outputted from the demand drivers model, distance and public transportation ride parameters require joint considerations for picking the best stations.

Based on the analysis of our results, we feel that the stations at the center of the city have reached close to their peaks. Of course, this is for the given level penetration and reach of the service in the population. With increase in advertising and further promotions, there is ample scope for improving the station performances. However in spite of the levels of saturation already reached in the city center, these locations are potentially more attractive than the locations in the outskirts. Thus the determination of new station locations would be an act of balancing between the less attractive outskirt communes versus the high yield saturated city communes. We would try to avoid the non-shaded and white shaded areas in Fig 23 to the extent possible. At the same time, we also aim to maximize the distance from the existing, upcoming and proposed stations when suggesting a new station.



Fig 23: Recommendations to locate future stations in Nice city

The first ten stations are all suggested from the city center which has exhibited the highest potential in our model. While these locations appear mathematically "attractive" by far

with respect to the other locations, it is also important to observe that these locations within the city are already facing saturation. Further addition of new stations could potentially cannibalize the performance of the existing stations. We believe that the model is underestimating the effect of interstation distances less than 300 m due to lack of modeling data. Locating new stations within 300-500 m of an existing station is fraught with risks. If the number of stations within a range of 500-700 m increases substantially, the impact on the performance of stations could be stark. But these locations within the city center have so much higher potential than the outer communes that these cannot be easily ignored.

The next dozen locations are usually suggested on the outer fringes of the city center. These locations may not perform as well as the locations within the city, but are nevertheless important in ensuring brand-awareness, options and connectivity to the Auto Bleue customers. These locations are usually chosen from bluish-green and white shaded regions on the map. As already mentioned, these locations are simply based on the mathematics and therefore lack intuition and ground realities. Thus the decision-maker is urged to apply their knowledge of the geography and landscape before selecting these locations for future stations.

Auto Bleue performance in the sub-urban areas of Nice has not been phenomenal, with the exception of the station at Cagnes-sur-mer. Application of our mathematical model on the given data actually shows that the sub-urban areas of Nice are indeed not as attractive to locate Auto Bleue stations as the city itself. Nevertheless, car sharing service is a social obligation and providing stations in far off locations would help improve the Auto Bleue brand in the long term.

It must be noted that the recommendations made for locating future Auto Bleue stations is also fraught with some risks, as mentioned below:

- One of the strongest assumptions made by us in relating the performance of Auto Bleue stations with locality based drivers is that the underlying relationship is linear in nature. While it makes some sense to use this assumption for a certain class of variables, linear regression is not certainly the state-of-the-art scientific method to explain station performances through these drivers.
- We have generally considered aggregated booking information to evaluate the performance of stations in the interest of time of study (three months). However it must be noted that the Auto Bleue business is fairly nascent and it would have been much better to do a fairly detailed analysis of the booking pattern progressively over time
- The consideration of performance of the existing system is based on the presumption that nothing else changes. We believe that Auto Bleue has already exhausted possibilities to place more stations in the "high performing heart" of the city. However a little more advertising of the service and reaching out to the potential customers can change the equations completely. If Auto Bleue can stimulate demand in this part of the city, there could be possibilities to locate more stations in this part of the city in the future.
- The model built for the sub-urban areas of the city is questionable due to lack of sufficient data. There are only five sub-urban quarters where Auto Bleue has a station and we also included Lindbergh in the study to increase the number of points. Mathematically, a model built with so few data points is expected to be unstable.
- The closest distance between two stations in the study has been St. François de Paul and Verdi with a distance of 250 m. We have made some assumptions regarding the impact of

station performances when the inter-distance between stations get further reduced. But these assumptions could not be validated due to lack of stable data. We believe that it will take some more months for the performance of station Auber and Gare Thiers to get stabilized, even though initial trends point to some fall in the performance of Gare Thiers.

6. Conclusions and Future Outlook

This study successfully presents some interesting insights into the optimization opportunity available in the car-sharing business. While most of the previous studies in the car-sharing context focus on the optimal sizing of fleets at the stations or employ crude clustering techniques to locate stations, our work presents a mathematical model to identify the n best locations based on the drivers of demand and diminishment of the drivers due to presence of multiple stations in the same catchment of the individual driver.

Our mathematical finds that the city of Nice has an attractive, but saturating, center which has the highest potential in terms of performance. As we move out of this heart, the service demand drops progressively. As a result, locating new stations is a complex trade-off between locating additional stations in the high-potential center versus untapped outskirts. The solution to our mathematical model points to the choice and selection of locations in both the heart of the city as well as the sub-urbs, indicating that the model has evaluated the trade-offs carefully.

One of the perceived risks for our model is that the inter-station distance effect is underestimated. In this context, we would advise Auto Bleue to commission new stations, especially in the high performing center, with great caution. We suggest that Auto Bleue place only a few stations in this zone at a time and then study the impact of cannibalization before commissioning new ones over subsequent phases.

This study can be used to not only determine the drivers of performance, but also improve the overall system usage and performance by analyzing target social and demographic groups for marketing campaigns. But it would be naïve to believe only on the results of our model. We strongly believe that locating new stations for Auto Bleue is as much an art as it is science. While our approach has primarily looked at the mathematics, it still needs to be backed up by local business knowledge and on-ground intuition. Though we must admit that there are some inherent risks associated with over-relying on the mathematical models, there is certainly merit in using the results of this study as the starting point for further exploration.

One of the main challenges that must be addressed by Auto Bleue and the public authorities in the near future is to design and operate the Auto Bleue service in such a way that it can complement the existing public transportation, without cannibalizing, while parallel improving the usage and performance of the car sharing service. In addition to optimizing the usage of Auto Bleue service, it must be aimed that the general population of the greater Nice area moves towards using public transportation. This can be addressed in a number of ways including greater understanding of the parking needs, better forecasting of origin-destination trips and improved pricing structures. Currently, it is felt that Auto Bleue performance is constrained by the fact that the users of the system are required to return the vehicle to the starting station. This is not a constraint while using the public transport. Though this must be validated through customer surveys, but there is a strong belief that Auto Bleue performance can be significantly improved by allowing customers to use the vehicles for one-way trips. However, allowing for one-way trips in the Auto Bleue system will bring about newer challenges. Some of these challenges are:

- Define pricing strategies to differentiate one-way trips from two-way trips, real-time booking from advanced booking, weekends from weekdays, high performing stations from low performing ones, and so on
- Forecast one-way demands between different localities
- Manage inventory of vehicles across different stations
- Resize the stations and the system

These could be some areas of study for a future research and would have immense social benefit for the city and sub-urban localities of Nice.

7. Acknowledgements

We are thankful to Veolia Transdev to sponsor this study. We are also immensely grateful to Mr. Mathieu Bernasconi and Mr. Chiocca Thierry for providing us information on the business and access to the data. Special thanks are also due to Mr. Jonathan Jouannet of Nice Cote d'Azur for providing us socio-demographic data for the city of Nice and its sub-urbs.

8. References

- Katzev, R. (2003). Car Sharing: A New Approach to Urban Transportation Problems. *Analysis of Social Issues and Public Policy*, 3(1), pp. 65–86
- Shaheen, S., D. Sperling, and C. Wagner (1998). Carsharing in Europe and North America: Past, Present and Future. *Transportation Quarterly*, 52(3), pp. 35-52
- Ciari, F., M. Balmer and KW Axhausen (2008). Concepts for a large scale car-sharing system: Modelling and evaluation with an agent-based approach, Arbeitsbericht Verkehrs und Raumplanung, *Technical Report 517, IVT, ETH Zürich*, Zürich
- Efthymiou, D., C. Antoniou, Y. Tyrinopoylos, and E. Mitsakis (2011). Spatial Exploration of Effective electric vehicle infrastructure location. *Transport Research Arena Europe 2012*, 0, pp. 1–10
- Uesugi, K., N. Mukai and T. Watanabe (2007) Optimization of Vehicle Assignment for Car Sharing System, *Lecture Notes in Computer Science*, 4693 (2), pp. 1105-1111
- Correia, G. H. A. and A. P. Antunes (2012). Optimization approach to depot location and trip selection in one-way carsharing systems. *Transportation Research Part E: Logistics and Transportation Review*, 48(1), pp. 233-247
- W.D. Fan, R.B. Machemehl, and N.E. Lownes (2008). Carsharing: Dynamic Decision-Making Problem for Vehicle Allocation. *Transportation Research Record: Journal of the Transportation Research Board*, 2063, pp. 97–104

Appendix



Fig 1A: Share of Managers and Experts across the city of Nice and its sub-urbs. Darker the shade, more the presence of Managers and Experts in that IRIS commune





Fig 2A: Share of Managers and Experts in the city of Nice

Fig 3A: Share of Car-ridership to workplace in the city of Nice. Darker the shade, more the car ridership



Fig 4A: Extent of Public Transport Ridership at IRIS commune level. Darker the shade, more is the public transport ridership



Fig 6A: Public Transportation ridership across East of Nice



