Mobility analysis by GPS data recorded on board of private cars

In the Italian vehicular fleet a percentage of about 1.5% is equipped with a GPS device for insurance purposes. The knowledge of the records collected by these vehicles, in the Florence province, during a period of one month, allows a traffic analysis on different municipal areas. For each municipality, the number of vehicles belonging to resident people and equipped with the GPS devices is estimated. This allows to normalize the calculated parameters. They are reported versus time and separated in resident and non-resident vehicles. Furthermore, the normalized parameters can be taken as indicators to assess traffic and migration behavior in a municipality as well as to compare different size municipalities

Giancarlo Giuli, Massimo Mancini, Paolo Zampetti

Analisi di mobilità da dati GPS registrati a bordo di autovetture

Nel parco veicolare italiano una percentuale di circa 1,5% è dotata di un dispositivo GPS. La conoscenza delle registrazioni rilevate in un mese da questi veicoli, nell'area comprendente la provincia di Firenze, permette di effettuare l'analisi del traffico nei Comuni nella zona. Per ogni Comune viene stimato il numero dei veicoli appartenenti ai residenti e dotati del dispositivo GPS. Questo consente di normalizzare le grandezze calcolate. Queste grandezze sono riportate in funzione del tempo e suddivise per veicoli residenti e non-residenti. Inoltre, grazie alla normalizzazione, queste grandezze possono essere prese come indicatori per valutare il traffico e i flussi migratori del Comune oltre a permettere un confronto tra Comuni di diverse dimensioni

After the diffusion of GPS technology, thanks to the quantity and quality of available data, new methods of traffic investigation are being developed, with many possible implications on policies for land management. This issue is the object of a project, PEGASUS, which involves ENEA and Octo Telematics[1]. In particular, Octo Telematics manages the data from all those vehicles that, for insurance purposes, have a GPS device (clear box) mounted on board. During the trip, with a frequency depending on the time elapsed (~30 s) or on the covered distance (~2 km), when required, the clear box records the UTC (Coordinated Universal Time[2]) time of the recording, the vehicle position, the instantaneous speed, the traveled distance from the previous recording and other parameters. The monitored vehicle population (hereinafter called "GPS vehicles") is fairly widespread throughout the country about 1.5% of the Italian private cars (2008) and is expected to increase.

Giancarlo Giuli, Paolo Zampetti
ENEA, Technical Unit for Advanced Technologies for Energy and Industry

Massimo Mancini
ENEA, Technical Unit for Energy Efficiency

In this article, some of the results obtained by using one month of tracking data are discussed. The analysis may be applied to a variety of applications: mobility planning and management, environmental impact assessment study, potential market definition of products and services, such as electric car[3], car sharing, innovative systems for local public transport, etc.
Within the project, the Octo Telematics Company has provided ENEA with data collected by GPS vehicles in the whole month of March 2008. The available data refer to the area of ~84x87 km^2, including Florence and Prato provinces, in addition to some towns of other provinces. With these data (~500˙000 records/day), detailed dynamics of traffic can be analyzed in different towns and compared with each other[4], [5], [6].

Normalization of results

A feature that makes these data very interesting is their almost uniform distribution in the whole Italian country. This is an important aspect because the values of the whole fleet may be estimated from the traffic data of the GPS vehicles.
The particular characteristic of the present study is the focalization of the traffic analysis on the municipal area rather than vehicles. Therefore, a normalization is needed and is obtained by dividing the values of the traffic parameters in a municipality by a characteristic number of the same municipality. In this work the characteristic number is identified in the number, N_R, of the GPS vehicles that belong to residing people in the municipality. This number choice is appropriate because this magnitude is proportional to the potential of mobility of the municipality itself and, at the same time, is independent from the traffic volume (influenced by internal and external factors of the municipal area).
In this normalization, the presence of other municipalities vehicles in the traffic does not cause any inconvenience.
At the municipal level, the usefulness of normalizing is evident when comparing, for example, the daily number of traveled kilometers by all GPS vehicles (veh-km), in Florence and Prato. It appears that Florence's values are higher, because of its larger population. This, however, does not necessarily imply that traffic is more intense in Florence than in Prato.
As regards the N_R value choice[5], let us examine the following case. A municipality with *only local traffic* is compared with another municipality characterized by the same N_R and the same local traffic, but, in addition, its traffic is produced by non-resident vehicles, with the same intensity of the local one, i.e., there is double amount of vehicles and of traveled veh-km. If, in each of the two municipalities, the total number of traveled veh-km is divided by the total number of vehicles, one obtains the average number of traveled km by a vehicle *(observation on the vehicle)*. The number is equal in both municipalities. This division does not show the different traffic level between the two municipalities. Otherwise if, in each of the 2 municipalities, the veh-km number is divided by the resident vehicles' number N_R *(observation on the municipal area)*, a double value compared to the first municipality is obtained for the second one. In this way, the effective difference between traffic levels is represented.
The graphs from analyses done on GPS vehicles are given on 3 different scales. The first one refers to data from GPS vehicles. The second scale refers to the normalized data, and reports the relative value (the value of the analyzed variable divided by N_R). With this scale it is possible to compare municipalities with different populations. The third scale refers to the whole fleet whose values are intended as an estimate of the real traffic in the considered area.

Determination of resident GPS vehicles' number in a municipality

The circulating or parked vehicles in a municipal area can be classified into *resident* and *non-resident*, depending on whether the vehicle owner resides in the municipality or not. In order to obtain an estimate of NR, number of resident vehicles, the available data are compared with the geographic reference information[7]. In this way, each registration is assigned to the municipal area in which it is detected. Then, by observing, for each vehicle and for each single day in one month, the municipality which the trip starts from at the beginning of the day, and which one the trip

finishes into at the end of the day, one can deduce the residence municipality. So it is possible to obtain the estimated number N_R of resident GPS vehicles.
In Figure 1, the monitored area is represented, bounded by the blue outline. In Table 1, the most populous municipalities included in the area are listed (col. 1) and the following corresponding values are reported: the number of registered vehicles[8] (col. 2), the estimated NR number of resident GPS vehicles (col. 3) and their percentage values (col. 4). In column 5 the population[9] is reported.
By considering the totality of the municipalities that are entirely included in the monitored area, we have 12,587 GPS vehicles (estimated), i.e., 1.33% of the total number of 945,958 registered vehicles. These total data are reported in the separate section on the bottom of Table 1.
The characteristic value N_R is used to normalize the results obtained for the considered municipality. The normalized magnitudes become *indicators* for a correct comparison of the traffic behavior between different municipalities.

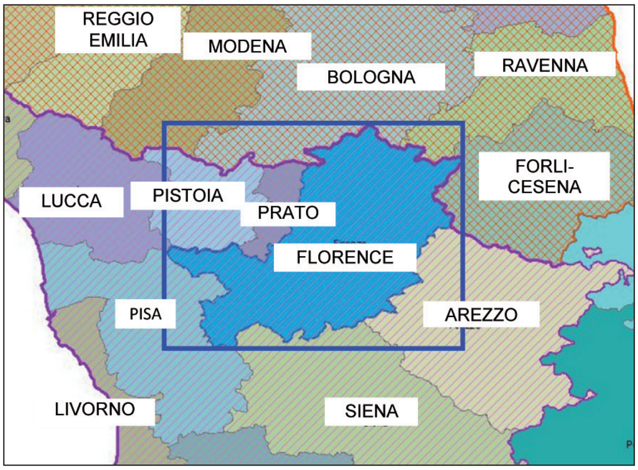

FIGURE 1 Detection area of GPS data (blue box)

Daily traffic intensity: vehicle presence and traveled distances

In order to get the estimated traffic intensity during the day, we can determine the number of circulating vehicles and the traveled distances. In the literature these two values are usually provided for time intervals more or less wide (1-4 hours). In this way the mean value in the interval is represented, without knowing the exact peak value and the instant at which it occurs. In this article the trend graphs are reported with continuous time scale, in order to better represent the rapid variations of the traffic.
In Figure 2, it is schematically shown how the continuous representation of the number of circulating vehicles is made. In the upper part, time histories of four vehicles are shown, represented in yellow during the parking time and in green during the circulating time. In the

Municipally	Regist. Vehicles[8]	Resident GPS Vehicles	% GPS Vehicles	Population 2008[9]
Florence	201,518	2,224	1.10	365,659
Prato	113,953	1,162	1.02	185,091
Pistoia	55,543	903	1.63	89,982
Empoli	28,891	813	2.81	47,181
Sesto Fiorentino	28,015	466	1.66	47,332
Campi Bisenzio	23,274	305	1.31	42,612
Bagno a Ripoli	15,687	206	1.31	25,885
Montecatini	12,976	133	1.02	21,156
Monsum. Terme	12,464	165	1.32	20,985
All Internal Municipalities	**945,958**	**12,587**	**1.33**	**1,622,997**

TABLE 1 Vehicular population in some municipalities

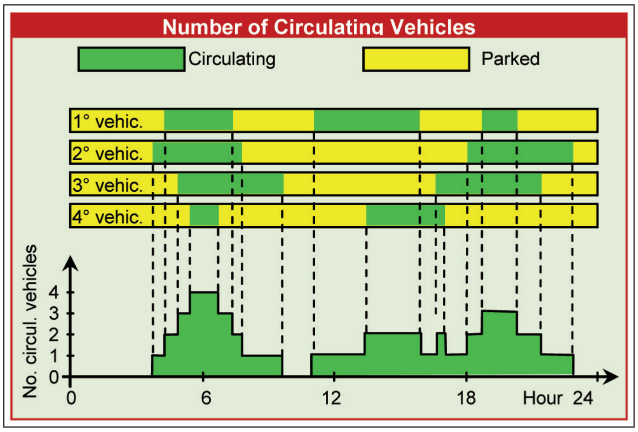

FIGURE 2 Schematic representation of circulating vehicles

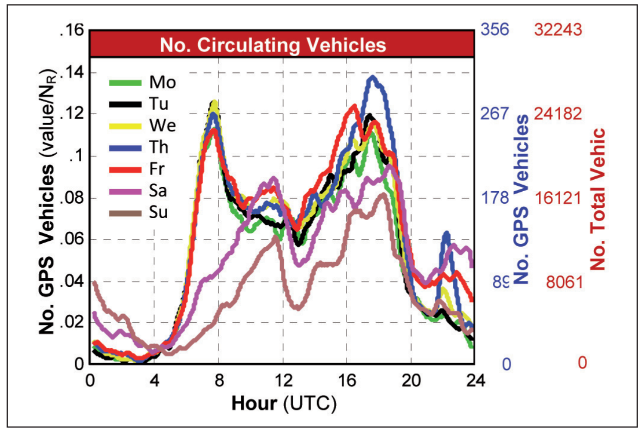

FIGURE 3 Number of circulating vehicles in Florence's municipality in the seven days of the week

lower part, the total number of circulating vehicles is reported moment by moment.

By applying this technique to vehicles in Florence's municipality, seven trends are obtained from the seven week days (3-9 March 2008), as shown in Figure 3, for the number of circulating vehicles. The time is reported on an hourly basis UTC (local time is 1 hour greater than UTC), as originally recorded.

The blue ordinate, shown at right, represents the presence amount of the GPS vehicles, whereas the number shown on the left ordinate represents the same data reported in a relative scale, i.e., divided by NR. Finally, in the third red scale on the right, the estimated real vehicle amount is shown starting from relative values. The graph shows:

- A perfect synchronization in the first 5 days of the week, at about 6-8 am and 7-8 pm.
- Less traffic on Saturdays and even less traffic on Sundays.
- The rush hour, at 8 in the morning, in the first five days of the week, moves to about 11:30 on Saturdays and Sundays.
- The traffic in the night between Saturdays and Sundays is 4 times higher than other days.

The black curve (Tuesday), shown in Figure 3 is reported in Figure 4. This curve is decomposed into two components: resident (green) and non-resident (red) vehicles. It can be noted that non-resident vehicles, are slightly more numerous than the internal ones up to 18.

To have a complete description of the municipal traffic, besides the number of circulating vehicles, the knowledge, in the municipal area, of the average speed or the total traveled distance per unit time by all the vehicles is required. These two quantities, as shown below, are dependent on each other. Here the second one is represented because it contains more information; for example, it allows an assessment of the pollution level.

Therefore, the total traveled distance by all GPS vehicles [veh-km] in the municipal area is calculated in a time window and is divided by the duration of this window. The quantity thus obtained is defined *total traveled* distance per *unit time* and is measured in [veh-km/h]. Because of its high variability over time, narrow time windows are used to have a practically instantaneous measure.

The *instantaneous* total traveled distance per unit time, on Tuesdays, is shown in Figure 5. As previously reported, three scales are shown along the ordinate axis. The relative value on the left is obtained dividing, by N_R, the value of the total traveled distance by GPS vehicles in the unit time. The total traveled distance, in a specified time interval, can be derived from the given trend, calculating the integral in the considered range. From the graph we see that in the rush hour, at 8.00 am, the total traveled distance for GPS vehicles is 5,560 veh-km/h. In fact, this distance is derived from the ~267 circulating vehicles (Figure 4), with

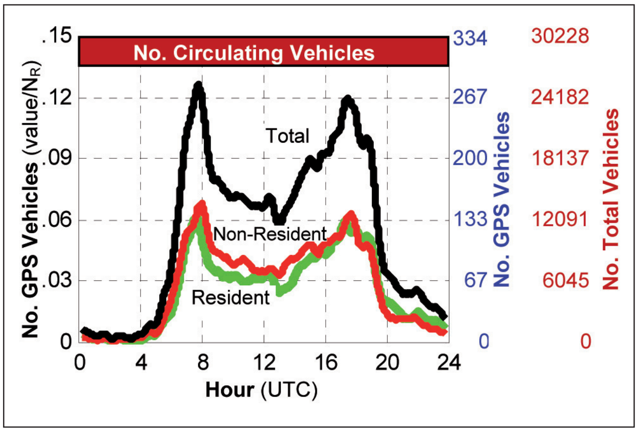

FIGURE 4 Number of circulating vehicles

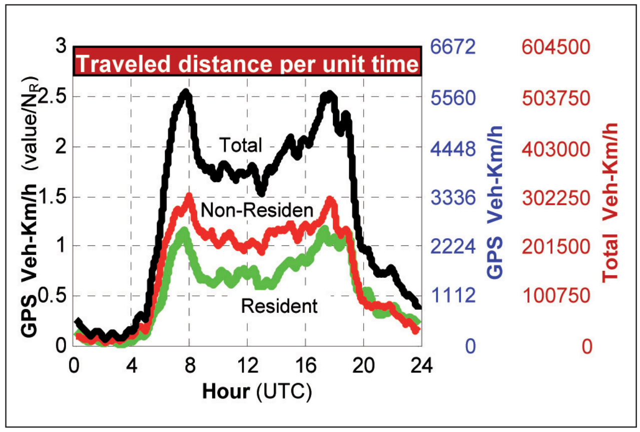

FIGURE 5 Traveled distance per unit time

an average speed of ~20 km/h, in that hour. Also in this hour the relative total traveled distance is 2.5 km/h, obtained by dividing the 5,560 veh-km/h for N_R (= 2,224) residents GPS vehicles. Then the relative distance has the following meaning: *the number of km that each one of the N_R resident vehicles should cover in one hour to produce the same traffic generated by the effectively circulating vehicles (resident and non-resident)*. For example, the relative value of 2.5 km/h, at 8 in the morning, corresponds to the distance that each of the 2,224 GPS vehicles, residing in Florence, should travel in one hour, at that time, to equate the value of 5,560 veh-km/h made by the 267 circulating vehicles.
It is interesting to note that the observed decrease, from 8 to 12 am, of the trend shown in Figure 5 is smaller than the one shown in Figure 4. This is due to the fact that in the middle of the day, because there is less traffic, the vehicles travel faster and therefore the mileage is proportionally greater than in the rush hours.
Another observation is that the red and green curves in Figure 4, between 7 and 18 hours, are almost coincident, whereas in Figure 5 the red curve, representing non-resident vehicles, is visibly above the green curve of resident vehicles. This is due to non-residents who, when entering the city, travel along stretches of highways with higher speed. The two trends, number of circulating vehicles and total traveled distance per unit time, complete the knowledge of traffic. In fact, also the instantaneous average traffic speed[6] is obtained from them, by dividing the values of the graph of Figure 5 by those of Figure 4.

Daily migration flows

An important issue in the study of mobility within a municipality is the assessment of migration flows, to and from a municipality, and the resulting change in the internal vehicular population, either circulating or parking.
Figure 6 shows, for each instant of the day, the variation in the number of vehicles with respect to midnight of the day before. Patterns refer to the seven days of the week taken in the Florence's municipality. The axis of ordinates, representing the variation of the number of vehicles shows, as previously, the three scales: the relative one on the left and the scales of the GPS (blue) and the total (red) vehicles on the right.
Figure 6 shows that at mid-morning, as first approximation, the estimation of the vehicular population increases with peaks of about 32,000 vehicles during the five working days. Otherwise on Sundays, at the same time, the vehicular population decreases by about 12,000 vehicles.
The number of commuter vehicles entering the town of Florence is higher than it appears in Figure 6. In fact, the variation of the vehicular population does not take into account that a large fraction of vehicles entering the municipality is compensated by local commuters leaving.

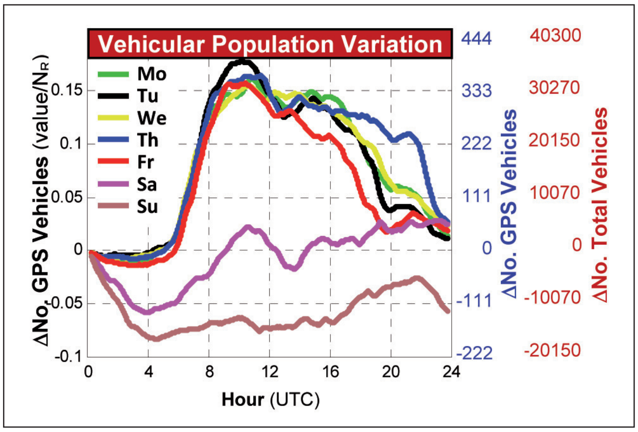

FIGURE 6 Vehicular population variation in Florence's municipality for 7 days a week

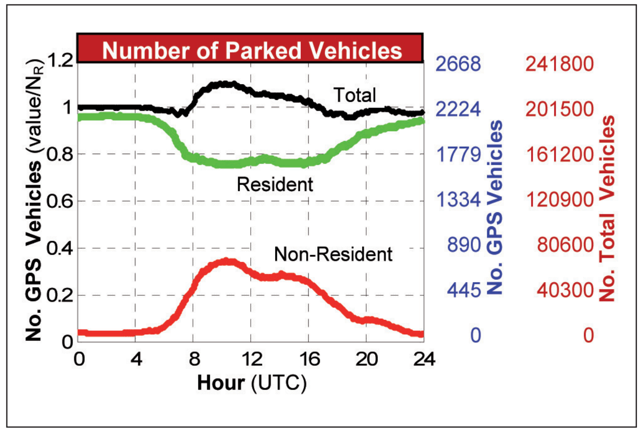

FIGURE 7 Number of parked vehicles in Florence's municipality

In order to study this aspect, the number of parked vehicles during the day can be observed, as reported in Figure 7, for Tuesday, on last March 3, distinguishing the parked resident vehicles (green) from those of non-residents (red). The total number of parked vehicles is shown in black. Again, three scales of ordinates are given: the relative one on the left, and the scales of the GPS (blue) and the total vehicles (red) on the right.
From the values of the green curve, read on the left scale, it follows that during the day the number of parked vehicles decreases by 25% for residents, with respect to the midnight hour, whereas it increases by about 35% for non-residents (red curve). This significant increase, given by non-residents, is partially visible looking at the total pattern (black curve), where the maximum value exceeds the midnight value only by 10%.
Since the time of parking is much longer than the circulation time, the shown trends of parked vehicles represent about 95% of all vehicles present in the municipality.

Comparison among different municipalities

As previously discussed, the data from GPS vehicles can be presented in a relative way. This opportunity is highlighted by comparing 4 different municipalities. The chosen municipalities are 3 provincial capitals (Florence, Prato, Pistoia) and Campi Bisenzio. The latter is located near a motorway junction. Changes in the number of vehicles are reported for Tuesday, March 4, 2008.

In the left graph of Figure 8, the trend of the number of circulating vehicles is shown, referring to the 4 municipalities, whereas in the right graph the total traveled distance for unit time is shown. In both graphs, the ordinates show the relative values.
Multiplying these values by the values of the third column of Table 1, the numbers of circulating GPS vehicle can be obtained. Similarly, one can determine the total number of circulating vehicles in a municipality, by multiplying the reported relative value by the corresponding value of the second column of Table 1.
The left graph of the number of circulating vehicles highlights that the 4 patterns shown are very similar. This is due to people needs, which are not dependent on the municipality. Unlike these trends, in the right graph, it can be observed that the total traveled distance in unit time depends on the considered municipality. In this case, the covered distance has a strong dependence on the road network. For example, Campi Bisenzio has a higher value of traveled distances in unit time than the other municipalities, because of its higher percentage of highways.
To understand the needs of mobility-related phenomena of commuting, a further comparison between the 4 municipalities is shown in Figure 9, where the change of vehicular population, compared to the midnight's, is represented.
It can be observed that, among the 4 municipalities, the growth of vehicles is higher in Florence and particularly

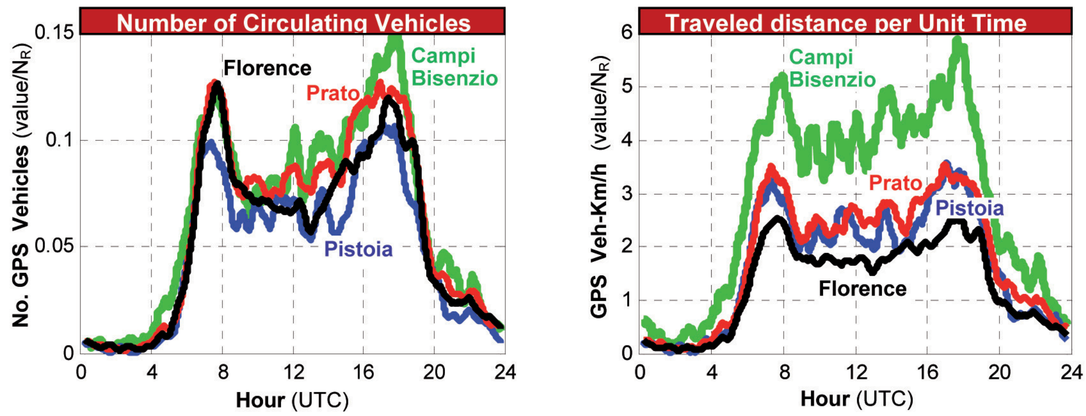

FIGURE 8 Comparison among 4 municipalities: number of circulating GPS vehicles (left) and traveled distance per unit time (right)

in the morning, because of activities related to offices, schools and universities.

The Campi Bisenzio's trends are quite variable. The presence of large shopping malls draws non-resident vehicles inside the city, mostly in late afternoon. Also a traffic component crossing the municipal area, due to the proximity of the highway, has to be taken into account. These movements produce high percentage variations on the municipal traffic because of the small resident population compared to that of crossing.

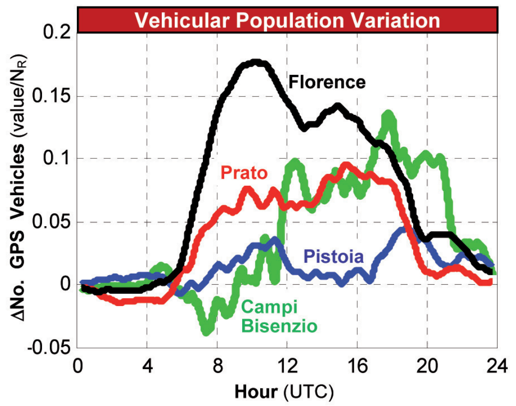

FIGURE 9 Change in vehicular population compared in 4 municipalities

Conclusion

The obtained results confirm that the proposed normalization makes the mobility parameters, obtained in different municipalities, comparable with one another.

Extending the GPS data processing, one can perform other analyses, in addition to those addressed, significant to get an accurate description of mobility. Some examples: presentation of results by isolating the component of resident, or non-resident vehicles in the municipality, instead of considering the total vehicular population; analysis of the vehicles occasionally entering the town (non-commuters), in order to quantify them into numbers and to see which days they are concentrated in and where they move; construction of matrices Origin/Destination of all vehicles, defining the departure/arrival points.

references

[1] OCTO Telematics Ltd., www.octotelematics.it

[2] http://en.wikipedia.org/wiki/Coordinated_Universal_Time

[3] G. Giuli, M. Mancini, G. Pede, Il mercato dell'auto elettrica, *Ingegneria dell'autoveicolo*, rivista dell'ATA, Vol. 64, n. 9/10, sett/ott 2011.

[4] G. Giuli, M. Mancini, P. Zampetti, Analisi della mobilità mediante dati GPS registrati a bordo di autoveicoli privati, *Mobility Lab*, n. 39, Mag/Giu 2011.

[5] G. Giuli, P. Zampetti, Analisi delle distanze percorse in un giorno da veicoli privati in un'area di traffico urbano ed extra-urbano, Technical Report ENEA, RT/2009/8/TER.

[6] G. Giuli, M. Mancini, P. Zampetti, Metodi di analisi dinamica della mobilità mediante elaborazioni di dati GPS, Technical Report ENEA, RT/2010/41/ENEA.

[7] CNEL-STATS, *Atlas of Infrastructure*, 2008.

[8] ACI (Automobile Club of Italy), *Autoritratto*, 2008.

[9] ISTAT (National Institute of Statistics).