Freight delivery services in urban areas: Monitoring accessibility from vehicle traces and road network modelling

Freight delivery services in urban areas: monitoring accessibility from vehicle traces and road network modelling

Abstract

Local Authorities plays a fundamental role in the management of city mobility and in accounting for the needs of different stakeholders involved in the urban freight transport. The aim of this study is to develop a method that could support the evaluation of the city accessibility for freight distribution services. As Local Authorities can use floating vehicle data (FVD), which are a current trend in mobility management, gaining new knowledge from data could be crucial to help the various stakeholders to better address their needs. Accessibility in urban areas is investigated through travel time estimations along the most frequently used routes connecting relevant nodes of the city and their average speed using a simplified road network model. After the description of the principal elements of the method, a test case is also presented for the urban area of Turin, Italy, to demonstrate the applicability of the procedures on a real scenario and dataset. The results confirm, also through the use of skim matrices, the value of FVD in assessing the accessibility of different zones interested in delivery operations, which may change over time, providing monitoring functions to urban logistics operators and Local Authorities in managing urban freight flows.

Keywords: urban freight logistics; urban accessibility; vehicle trajectories; traffic congestion; transport modelling; Local Authorities

1. Introduction

1.1. Urban freight distribution scenario and planning tools

Freight distribution is increasing its role in urban road traffic, owing to the growth of internet shopping, which is partially substituting traditional goods purchasing. In 2016, the e-commerce market accounted for 19.6 billion € in Italy, an 18% increase compared to 2015 (Freight Leaders Council, 2017), whilst the same market was 530 billion € for Europe in 2017, which was 15% higher than the previous year (European Ecommerce Report, 2017). As a result, an increasing amount of goods is travelling within cities and is delivered directly to individual consumers instead of arriving in bulk to select store locations. Obviously, this reflects the pattern of urban transport demand as an addition to the traditional distribution to shops.

At the city level, the greater number of vehicles travelling around a city to make such deliveries adds to the existing traffic characterising an already congested road network. At the environmental level, recent analyses reveal that "in Europe urban freight is responsible for 25% of urban transport related CO_2 emissions and 30 to 50% of other transport related pollutants" (Meyer and Meyer, 2013,p. 4). Consequently, public authorities are being asked to propose and test policies to control and manage traffic in cities with the expected aim of reducing air pollution, as well as to protect historical centres and monitor land use. The methodology proposed in this paper, based on the availability of a floating vehicle data (FVD) dataset, can be a possible tool also for Local Authorities to evaluate possible effects of those actions easily reversible, for which a continuous monitoring procedure could measure the impacts directly on the field. This approach is more simply implemented if compared with other model-based approaches, which often requiring simulations. Such measures can be managed in the framework of Sustainable Urban Mobility Plans (SUMPs), which are medium-term planning tools that are becoming mandatory for cities and metropolitan areas in European countries (ELTISplus, 2017). The evaluation of such policies and their effects on citizens and stakeholders should be assessed in both current and alternative scenarios by specific and measurable indicators. This monitoring of the planning process and of the impact of the implemented measures on city mobility is a fundamental requirement to promote actions which effectively contribute to achieving the expected benefits (Ambrosini et al., 2010). In addition, for the wider movement of goods travelling within cities, local authorities are required to propose specific regulation strategies for urban freight distribution (Kiba-Janiak, 2017). In fact, the European Union requires cities to define urban freight plans to study measures to modify the efficiency of urban logistics, with the challenging objective of reducing the related greenhouse gas emissions and noise (Fossheim & Andersen, 2017). More specifically, SUMPs must include specific actions in their logistic components for the so-called Sustainable Urban Logistics Plan (SULP) (Ambrosino et al., 2015).

As confirmed in a recent aforementioned study, deliveries have a significant impact in terms of traffic congestions around a city because they account for approximately 10–15% of all urban kilometres travelled (CIVITAS WIKI consortium, 2015). More specifically, approximately 25–30% of urban deliveries are carried out by light vans in Europe (ALICE & ERTRAC, 2015). Hence, information extracted from such rich and wide datasets could provide actual feedback on freight traffic trends through the definition of specific indicators. Moreover, proper analysis could help in assessing the impact of the measures proposed at the city level, for example in the SULP (Diana et al., 2020).

1.2. Research contributions on network monitoring and accessibility

The aim of this study is to develop a method to measure city accessibility for freight distribution services using the positioning data collected during the trips taken in the van.

In general, their travel time and the average speed can be easily estimated along the most frequently used paths that connect relevant zones in a city if vehicle data are frequently recorded and integrated with a geographic information system (GIS) (Pascale et al., 2015; Taylor et al., 2000). For example, Greaves and Figliozzi (2008) installed commercial global positioning system (GPS) devices in selected vehicles participating in a travel survey to detect the freight tour features in cities, and were able to record second-by-second trip data during delivery operations for the period of the experiment. A similar approach was applied by Ben-Akiva et al. (2016), in which GPS loggers were fitted in participants' trucks and integrated with a web-based survey to detect route choice behaviour. As an alternative, smartphones can be used to detect high resolution vehicle traces (Ge & Fukuda, 2016), as demonstrated by Gonzalez-Feliu et al. (2013), in which data were analysed at the route level with primary focus on delivery bays. No mapping was performed to a road network model. In another study, Yang et al. (2014) used second-by-second GPS data to identify urban freight delivery stops. However, in practical applications, data available from commercial services (taxi or freight vehicle fleets) may not be collected with such high sample rates, and in these cases, alternate techniques need to be applied to extract useful information.

Cui et al. (2016) applied a method to estimate city accessibility by using GPS data collected from a taxi service, in which the data sampling ranges were from 30 s to 2 min, which did not achieve the 1 Hz sampling rate required to reliably map traffic conditions along road links. In their study, a network model is not used and travel times are related to points on a map where positioning data are available for vehicles. Then, trips are built to estimate accessibility using a zone-based approach considering the points belonging to zones as starting or ending points. Other studies in the literature use GPS data for different applications, for instance Sharman and Roorda (2013) used GPS truck data to study the inter-arrival duration, defined as the time between arrivals at a destination of two successive vehicles operated by the same carrier. Hess et al. (2015) proposed a novel application in route choice modelling using GPS data focussing on heavy goods vehicles. With respect to traditional survey techniques, GPS data provide a better estimation of parameters such as route length, number of stops, fuel consumption and CO_2 emissions (Pluvinet et al., 2012).

GPS data sent with a random sampling rate by transponders (from 1 to 8 pings/h) on trucks were used to analyse the functional corridors of the state of Mississippi, USA by Holt et al. (2017). They collected more than 26 million individual truck data points over four years, which were mapped directly on the links of a GIS road network to assess its performance in predicting freight transport statistics, such as travel time, average speed, and congestion in relevant connections.

The amount of GPS data needed to provide an accurate and time-dependent speed estimation for real-time applications along selected corridors was investigated by Patire et al. (2015) for different sampling and penetration rates and for comparison with other technologies (e.g. inductive loops, Bluetooth). The study found that even though a higher average sampling rate produces more data points, it is preferable to collect data from different devices to improve the accuracy of the travel time measure on roads. Therefore, a higher penetration rate is more effective than a higher-resolution rate. For this reason, the approach developed in the present study uses an available dataset with a low sampling rate, and relies on the detection of multiple vehicles at the same node of a network model.

1.3. Exploitation of the method

As will be explained in more detail, the proposed method has a twofold relevance at the urban level. In fact, it can be exploited by public authorities to analyse the current network performance regarding freight delivery and to plan future measures (e.g. the introduction of a booking mechanism for loading/unloading bays, a special policy to dynamically manage access to restricted traffic areas, realisation of a freight route planner app to optimise deliveries (Pronello et al., 2017)), but also by delivery service operators for shifting delivery times from congested to off-peak periods. Policies that shift urban goods deliveries from daytime to off-peak hours have the potential to increase the efficiency of freight distribution and reduce negative external impacts. The interaction between public authorities and delivery service operators, sharing the monitoring approach proposed, can also lead to the redefinition of the policies, including for example the rules of restricted traffic zones (time slots, access rights, fares) or even the use of dedicated lanes for the exclusive use of commercial vehicles at certain times and along selected routes. Freight transport management in urban areas can then be based on the observed traffic conditions. For instance, Fu and Jenelius (2017) used vehicle GPS probe data, fleet management data, and logistic information to assess the impacts of specific policies in Stockholm, Norway. According to de Palma and Lindsey (2011) various congestion pricing schemes can be adopted in urban areas to reduce congestion. The traditional approaches are based on facilities use, on cordoning-off crossings to protect specific areas of the city, or on zonal pricing to modify the behaviour of freight distributors. However, considering available technologies, such as a global navigation satellite system (GNSS), more advanced schemes can be applied, such as distance-based pricing specifically set for vehicle types and time of day. For example, the Off-Hours Delivery experiment in New York City, USA, with a time-of-day pricing strategy shifted only 20% of carriers, but the savings in terms of travel time to all highway users was approximately 3 to 5 min per trip and to carriers that switched to off-hours was approximately 48 min per delivery tour, with economics savings estimated between 100$ and 200$ million per year in travel time saving and pollution reduction (Meyer & Meyer, 2013).

The design process involving any of these actions needs knowledge of the network conditions and measurement of the accessibility to urban zones for delivery operations, which should be monitored during the implementation phase to adjust, if necessary, the tolling scheme. To achieve this objective, the raw data chosen for the methodology presented in the current study are van GPS traces because of their targeted information value for freight delivery in urban areas and their easy availability. Indeed, they are commonly exploited to monitor vehicle routes and to record stops for loading, unloading, and parking (Pirra & Diana, 2019). Good cooperation between the data owners (operators) and Local Authorities is fundamental because this allow the collection of such information. Therefore, Municipalities may establish some long-term agreements with those operators, such as special area access permission and operational licenses in exchange for the provision of that kind of datasets.

2. Methodology

In this paper's framework, "accessibility" is defined as the ease and extent to which road networks enable delivery vehicle fleets to reach the various zones of a city. On the whole, a variety of methods have been developed for measuring accessibility and they can be classified according to Geurs and van Wee (2004) as the following:

- Infrastructure-based measures, which analyse the performance of a transportation infrastructure.
- Location-based measures based on indicators related to the spatial distribution of activities.
- Person-based measures at the individual level, considering individual requirements and limitations.
- Utility-based measures, which consider the benefits that people derive from levels of access based on spatially distributed activities.

Additional categories are provided in Curl et al. (2011):

- Cumulative measures, which represent the accessibility at a location to another or set of destinations.
- Gravity-based measures, a weighted extension of cumulative measures.

The approach chosen for the present study can be categorised as a mixed approach, because the measures used, such as travel times on the road network, describe the function of the transport system (infrastructure-based). Additionally, accessibility is defined as the degree to which two zones in the study area are connected (location-based) by using the travel time and speed of a set of vehicles, estimated by their positioning data.

The main steps of the procedure, shown in the scheme of Fig. 1, and also applied in a case study described in Section 3, can be summarised as three main steps.

2.1. Construction of the a priori *network*

The first step of the methodology requires the creation of a sketch model of the road network, called *a priori* network, in which main links, nodes, and centroids are identified and classified on a georeferenced map. In this phase, a simplified links classification may be applied, identifying the motorways and the main roads in the city using a traffic modelling tool to create this high-level representation. Node selection can be performed by considering all intersections of the urban motorways regarding their connection role in the road structure, whereas only a subset of the urban area intersections should be selected based on their relevance to routes connecting the different zones of a city. Local and secondary streets should not be included in this simplified road network model. Indeed, the focus on the zones accessibility requires to consider only the main roads that could be followed by the vehicles during their travelling around the city for their deliveries.

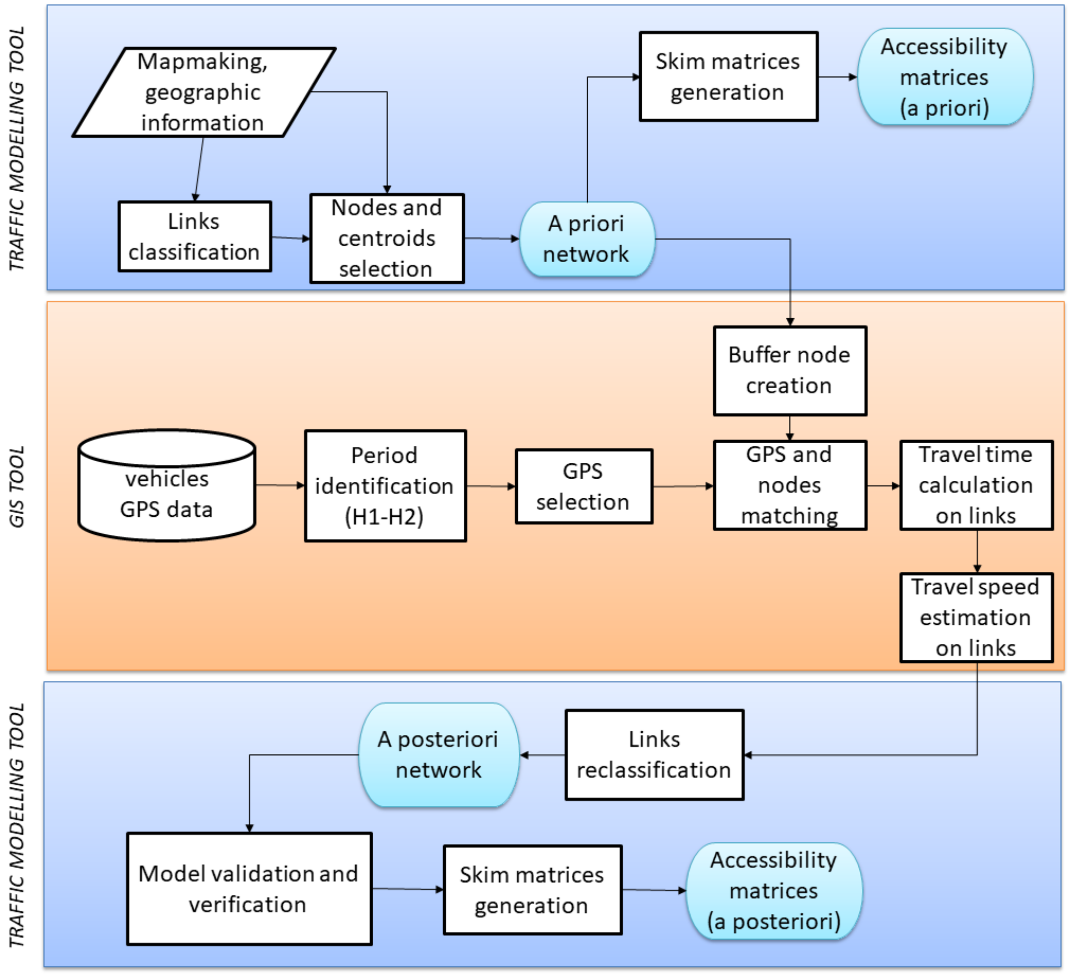


Fig. 1 Flow diagram of developed methodology and data used in the three main steps

2.2. Travel time estimation from GPS data

Travel time estimation from GPS data is performed for the links connecting the selected nodes by exploiting the positioning data collected by light vans during their usual delivery operations in the city, which can be provided by tracking and tracing systems already available. The travel time calculation aims to better define the network features and road types for homogeneous time periods over a day. In this phase, a buffer for each node type is also defined to effectively detect as many vehicles as possible at road intersections and consequently their travel times along links. Each node of the *a priori* network and its buffer area can then be used to detect the moment when an equipped vehicle crosses the related road intersection. Mapping the vehicles at the nodes rather than along the arcs increases the chance of detecting them in the case of low sampling rates, because they usually spend more time at intersections (in particular when the traffic signal is red), and because the positioning accuracy based on a satellite's line of sight is generally higher. This operation can be implemented through the creation of a circular boundary area around each node of the network by using GIS software. To extract information, only the vehicle positions in those areas are taken into account, storing the vehicle identifier, date and time of passage, position (latitude, longitude), course and speed. An explorative analysis can be conducted to determine the proper dimension(s) of the boundaries in terms of their diameter to ascertain if the number of detected vehicles can be increased whilst taking into account the quality of vehicle positioning at intersections. Because the accuracy of GPS positioning is reduced, especially in urban canyons, a proper boundary around each node is determined to increase the chance of detecting vehicles at each node, as explained in the application example in Section 3.

The travel time along links is estimated without applying classic map matching procedures based on a link approach, as in Holt et al. (2017), in which the vehicle position is associated to the links. Matching is accomplished by focussing on the time when a vehicle is detected at selected nodes of the road network. The time of detection is then referred to each vehicle travelling around the city on a specific working day (of the monthly observation period). A further key issue is to determine whether a vehicle has actually moved along the arc in its travel between the two nodes at the arc's extremes by comparing the course of each GPS recording, i.e. the driving direction of the vehicle, with the direction of the arc. More precisely, all the courses associated with GPS recordings of the selected vehicle in its travel between the two nodes are considered to verify that they are similar to the corresponding value of the arc bearing. Some variation in the course values is tolerated in the algorithm to include any measures that are different from the arc bearing value simply because of the road's curvature. More precisely, the root mean square error between all GPS courses and the bearing is computed: if it is less than 50 degrees the vehicle is assumed travelling through the arc under consideration and without deviations(Pirra & Diana, 2019).

Then, estimation of the link travel time (TT) is derived by computing the difference between the timestamps of the first recording in the boundary around the origin node and the first recording registered in the boundary around the end node. This value is related to phases when the vehicle is moving and those when it stops due to traffic conditions (delays at intersections, congestion, traffic lights) or service operation (e.g. the time required to make a delivery). The overall time interval for a series of subsequent 0 speed recordings along the selected arc is calculated and its duration is named ST (stop time).It is possible to find various values (Ns) of ST for each link, namely ST_i, i= 1,…,Ns. It is thus necessary to remove them to obtain the "real" travel time along the arc (TT_r). However, while dealing with congestion, we could think to consider the stops due to traffic condition as part of the time required to travel along a road. Therefore, only the ST_i associated with the deliveries has to be removed from the travel time TT. A specific threshold of 120 s is defined to differentiate these two cases. Time ranges ST_i shorter than this value are commonly associated to typical maximum duration of a stop for yielding or at traffic lights, whereas service stops are normally longer (S. Greaves and Figliozzi, 2008). Thus, when the computed ST_i is longer than 120 s, it is considered as a service stop and has to be removed from TT, otherwise it could be considered as part of the time necessary to travel along the road. The final value TT_r for each arc is thus obtained as

$$TT_r = TT - \sum_{i=1}^{Ns} ST_i$$

Where Ns is the total number of stops intervals found and

$$ST_i = \begin{cases} 0 \ \ if \ ST_i \leq 120 \ s \\ ST_i \ if \ ST_i > 120 \ s \end{cases}$$

2.3. *Construction and validation of the* a posteriori *network*

The data derived from the previous steps of the methodology are exploited to define the final network, called *a posteriori*, which represents an updated model with estimated travel time information and a more realistic road classification based on observed travel speed. In fact, the main street characteristics originally associated in the *a priori* network are refined using the travel time information extracted from the GPS traces dataset. Moreover, the known link lengths along with the estimated time necessary to travel along each of the arcs of the network are used to compute average speeds, thus creating a broader and more reliable classification of the links. Many values of travel time can be associated to a certain arc during the investigation period (one month). The speed value used to refine the links classification is therefore computed starting from the average travel time obtained by removing the outliers to reduce the influence of exogenous factors, such as road work, that could worsen traffic conditions on some days of the observation period. To improve consistency, classes can be defined based on the average speed distributions of mapped links presenting at least 10 measures (after outlier removal) and the shape of the plot, as it will be clarified the case study in Section 3.3.

At this point, model verification is necessary to ascertain if the travel time values estimated to measure the accessibility among selected zones provide consistent values compared to those supplied by map providers on the web (e.g. Open Street Map, Google Maps, Here). Moreover, a validation of link classifications is performed to check if the simplified approach used yields acceptable results for the estimation of accessibility. In fact, as explained previously, each link is assigned to a specific class according to the average speed derived from the previous step of the methodology. This new categorised value is associated with each of the links. This is an approximation that allows better management of the model and guarantees negligible loss of information with respect to the travel time estimated between zones. As an alternative approach, the specific speed values estimated for each link can be used to map the accessibility to the zones of the study area. Therefore, the validation process involves a comparison between these two scenarios to validate the approach and the classification adopted.

A further step requires investigating the accessibility matrix estimation for the zones of the study area by considering skim matrices of travel times along the best route generated by the traffic modelling tool for the *a posteriori* network at two principal time periods of each day and comparing them to similar results from the *a priori* network. A skim matrix includes impedances between zones and can provide numerical quantification on the accessibility of different parts of the study area (Mcnally, 2007). Therefore, it is exploited to evaluate the city's accessibility by considering the travel time (min) and distance (km) indicators. The analysis, performed using the OmniTRANS tool, focusses on the computation of the shortest path between the various centroid pairs, where the algorithm considers the travel time or the distance as the main link parameter.

3. Application to a case study

The proposed methodology is applied to a real case study represented by the city of Turin, capital of the Piedmont region in north-west Italy. Its centre includes more than 10,000 economic activities. The wide diffusion of e-commerce deliveries coupled with normal freight transport represents 8% of Turin's total traffic (Freight Leaders Council, 2017), with the associated need for proper accessibility evaluation. Due to the interest on the topic, the city has been actively involved in European projects dealing with urban freight mobility. Moreover, Turin has implemented a set of 'push and pull' measures combining both incentives and restrictions for those operators that follow a Freight Quality Partnership Agreement in their delivery activities. Most of these measures aim at reducing and rationalising deliveries in the city centre, which is characterised by a limited traffic zone.

3.1. Construction of the a priori network

A preliminary network has been imported from OpenStreetMap to map the accessibility in the selected study area (Fig. 2). This network contains a large number of arcs and nodes (in our case, more than 10,000 links and 5000 nodes), in which the network connectivity is not always guaranteed and includes some link directions that need to be checked. For this reason, as described in Section 2, an *a priori* network was created manually based on this georeferenced map by selecting principal nodes and links.

The network in our case study contains 408 two-way links, including 84 connectors, 110 nodes and 18 centroids. Only two main types of links are defined to simplify the network:

- "Motorway" includes the links for urban motorways. The speed setting is 80 km/h according to the authors' experience of the average speed during congested periods.
- "Road2lanes" includes all other links. The speed setting is set to 30 km/h (although the maximum speed is 50 km/h) to consider the presence and effect of secondary intersections along the links affecting traffic conditions.

One internal centroid is located at the Turin city centre, whereas 17 external centroids are chosen according to their relevance in terms of connections with the urban network, including the main high-speed road (A55 Turin Ring Road), for its relevance to freight distribution vehicles (Fig. 3).

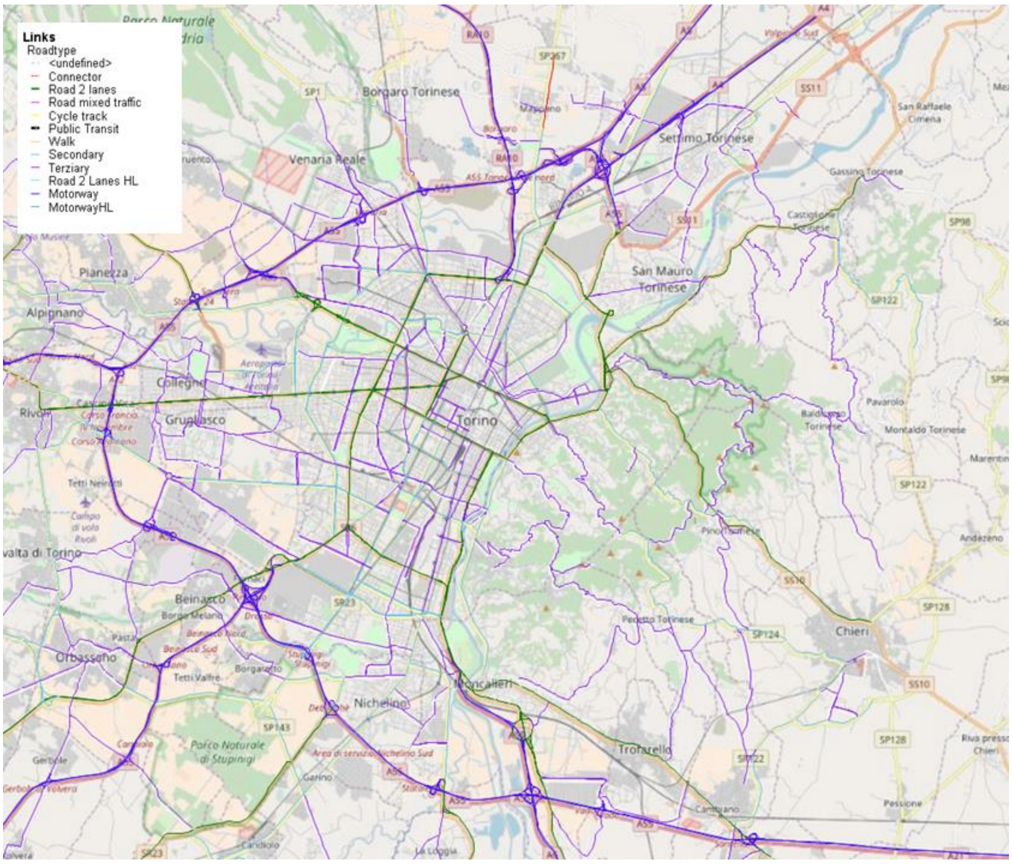
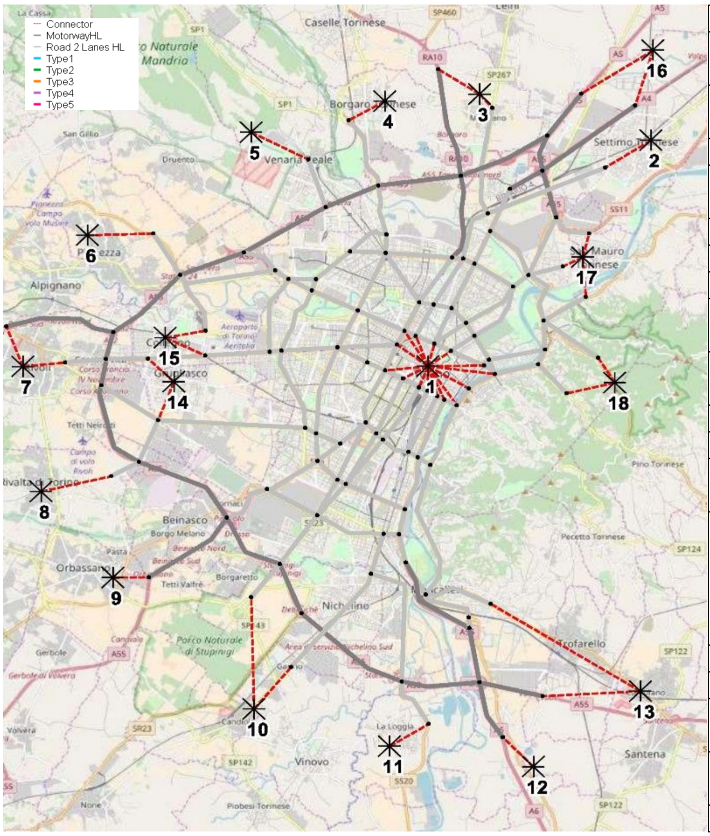


Fig. 2 Original network of Turin area (Source: OpenStreetMap)



Centroids ID	Name
1	Turin City Center
2	Settimo Torinese
3	Mappano
4	Borgaro Torinese
5	Venaria Reale
6	Pianezza
7	Rivoli
8	Rivalta di Torino
9	Orbassano-Sito
10	Candiolo
11	La Loggia
12	Highway South (A6)
13	Santena-Trofarello-Cambiano-Moncalieri Chieri
14	Grugliasco
15	Collegno
16	Highway North (A5-A4)
17	San Mauro-Pescarito
18	Chieri

Fig. 3 *A priori* network of Turin area: Motorways in dark grey, Road2lanes in light grey and connectors in dashed red (Source: OmniTRANS model)

3.2. Travel time estimation from GPS data

The proposed method is applied on a dataset consisting of 360,820 GPS positions in Turin related to vehicles (light vans) belonging to logistics fleets delivering goods throughout the city (Pirra & Diana, 2019). More precisely, GPS traces were collected for 28 different vans in the period from April 29 to May 29, 2017, however only 23 vehicles were detected while travelling within the selected area during work days. Each recording includes time and day, latitude and longitude, instantaneous velocity and bearing.

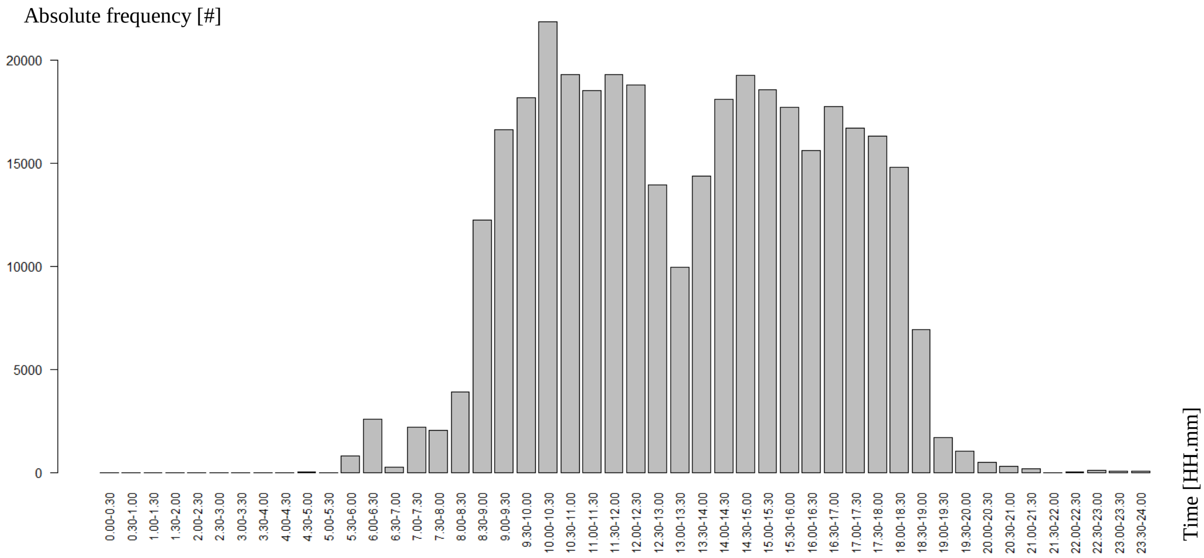


Fig. 4 Recording distributions according to the time of day for 30 min time ranges

The time periods investigated are not referred to an hourly basis; their range is selected according to the frequency of the position data collected at various moments of a day (Fig. 4). Considering the specific characteristics of the dataset (delivery operations, small number of vehicles), wide time periods must be set to capture a larger number of vehicles and to refer the estimated speed to homogeneous periods. Moreover, it is necessary to recall that the vans are travelling around the city to deliver goods to shops mainly during normal business hours. Off peak travel times are not considered in the analysis, because the influence of congestion on the travel could is not relevant. According to such observations, the analysis in subsequent sections is applied based on the following two-time ranges:

- H1 → 9.00 - 12.30 a.m.
- H2 → 4.00 - 6.00 p.m.

As discussed in the Methodology section, it is necessary to detect as many vehicles at road intersections as possible to compute their travel times along links. Two main classes of nodes are identified to adapt a boundary area to the relevance of the road intersection: one represents the case where a “Motorway” road is present (meaning Motorway to Motorway and Motorway to Road2Lanes), while the other includes the crossings of “Road2lanes”. For both classes, six radii were evaluated, from a minimum of 50 m to a maximum of 300 m, considered in 50 m increments.

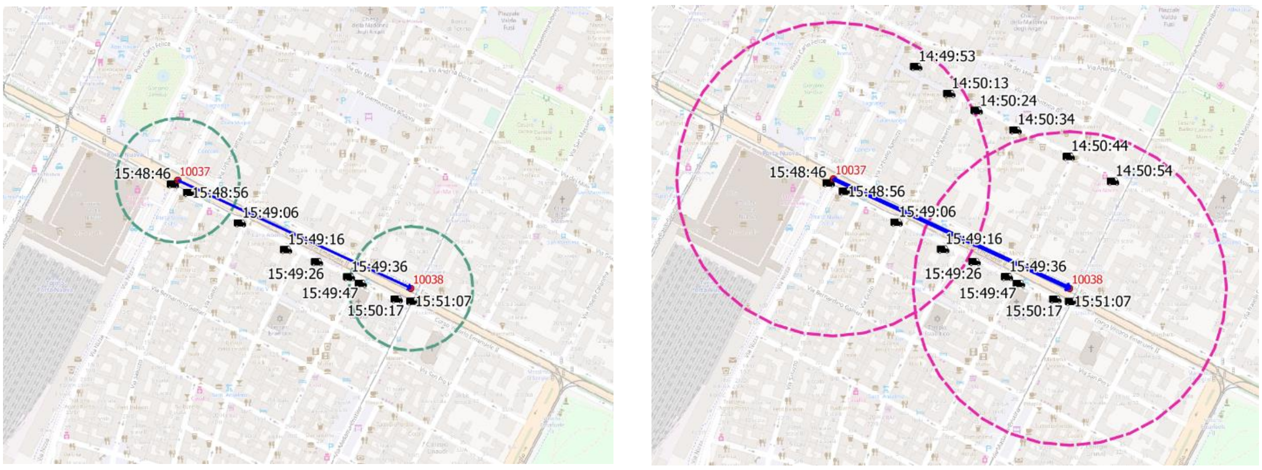
The selected values of the radii for the different classes of nodes are given below and Fig. 5 shows examples of two common node types.

- Intersection of two “Motorway” roads or “Motorway” to “Road2Lanes” → r = 200 m
- Intersection of two “Road2lanes” roads → r = 100 m

(a) (b)

Fig. 5 Examples of the two classes of nodes and positioning data: (a) intersection connecting two “Road2lanes” and (b) a node at the crossing of a “Motorway” and a “Road2lanes” (Source: QGIS).

These values are selected by combining a numerical analysis with evaluation of the map. In fact, the number of links (roads) where vehicles have travelled is computed for the various combinations of radius values. As expected, higher numbers of passages are detected if the boundaries are wider for both classes of nodes (e.g. 250 m). However, some problems regarding the quality of the results could arise in those cases. Fig. 6 shows an example for a link in the city centre belonging to the “Road2lanes” class and the connecting nodes 10037 and 10038. Two different radii are proposed for the boundaries, namely 100 m (Fig. 6(a)) and 250 m (Fig. 6 (b)), as well as a selection of positions recorded for two vehicles travelling in that area of the city. Fig. 6(a) shows that a vehicle has effectively travelled along the selected arc because it has been localised in the 100 m boundaries around both nodes. By contrast, the image presented in Fig. 6(b) highlights the role of a proper radius. In this figure, the radius is set too high, and other vehicles travelling along parallel roads can be erroneously taken into consideration. To avoid this drawback, the selected radii are those aforementioned.



(a) (b)

Fig. 6 Example of a link (blue line) connecting nodes 10037 and 10038, both representing intersections of two “Road2lanes” in the city centre. The images represent two possible radii length: (a) 100 m and (b) 250 m (Source: QGIS).

As discussed previously, the procedure is conducted in two steps by dividing the recordings according to the time ranges H1 and H2. Considering the small number of vehicles included in the dataset, the application of the methodology described in the previous sections provides observed travel time only for a certain number of arcs in the *a priori* network. As indicated previously, GPS traces were collected for only 23 vehicles, with 22 of them found travelling along the *a priori* network arcs. However, a detailed count yielded values of travel time for 216 of the 324 arcs composing the network in the time range H1, whereas this number declined to 155 in H2. As explained previously, the dataset encompassed one month of recordings, thus, each link could have been travelled more than once in each time interval. For example, it is possible to find up to 63 values for the same arc travelled by different vehicles. This is logical because, in the cited case, the corresponding road is one of the main access routes for vehicles entering the city from the north, where some of the main logistic structures are located. However, because a single and representative value of speed is associated with each link during a given time period, this can be estimated by considering the average travel time. A further refinement is actually proposed to improve the reliability of the final value obtained; the average is computed only after removal of the outliers from all possible travel times found for each specific arc. By applying this operation, it is possible to avoid the influence that unexpected fluctuations in the values collected could have on the final average travel time. On the whole, outliers were found and removed from 21% of the arcs in H1 and 17% in H2.

3.3. Construction of the a posteriori *network*

As explained in Section 2, the information derived from the GPS traces dataset is exploited to compute the average speed for each a priori network link, which is determined from the relationships among the distances between nodes and the corresponding average travel times. In such manner, the original road characteristics associated by default to the various arcs are now closer to reality, as perceived by vehicles travelling within the city. To gain consistency in the classification, this operation is performed considering only those arcs with at least 10 values of computed speed after outlier removal for H1, namely 18% of all arcs (38 of 216). The average length of these links is 1.7 km with 89% of them longer than 500 m.

Fig. 7 shows the approach adopted to create new classifications for the *a posteriori* network. The average speed values for the 38 arcs are firstly organized in decreasing trend (red line), based on the corresponding minimum and maximum values (green and blue lines). Then, five new classes are defined from the distribution of values and slopes in the plot. In particular, the limits of Type3 have been identified according to the highest slope variations, and then, two additional types for higher speeds and two additional types for lower speeds are introduced, approximating the shape of the average distribution. The minimum and maximum distributions confirm that the range around the average is quite narrow, with some exceptions, which pertain to short links that have a negligible effect on the travel time estimation along the routes. Considering the larger size of the GPS traces dataset recorded in time period H1, this period is used as the reference for class definition.

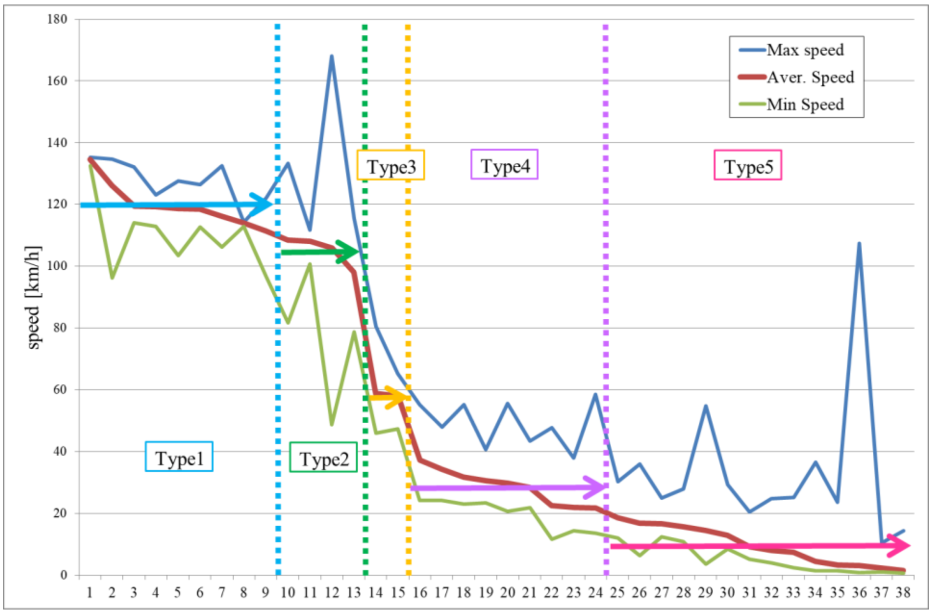


Fig. 7 Average speeds ([km/h) for the 38 arcs of the *a posteriori* network used to define the new road classifications

Table 1 provides further details on these new classes: their names (first column), the extreme values of average speed used to assign each arc to the different classes (second column) and the corresponding average travel speeds that are associated with each road type (third column). Moreover, the number of links of the *a posteriori* network that are currently assigned to each of the five classes is provided, both for those arcs with at least 10 measures (fourth column) and for those with at least 5 measures (last column). Note that the majority of arcs fall in the “slowest” class. For a deeper investigation, it could be useful to check where the different types of links are located on the city map to gain profitable information on how the logistic fleet “perceives” accessibility and mobility around the city.

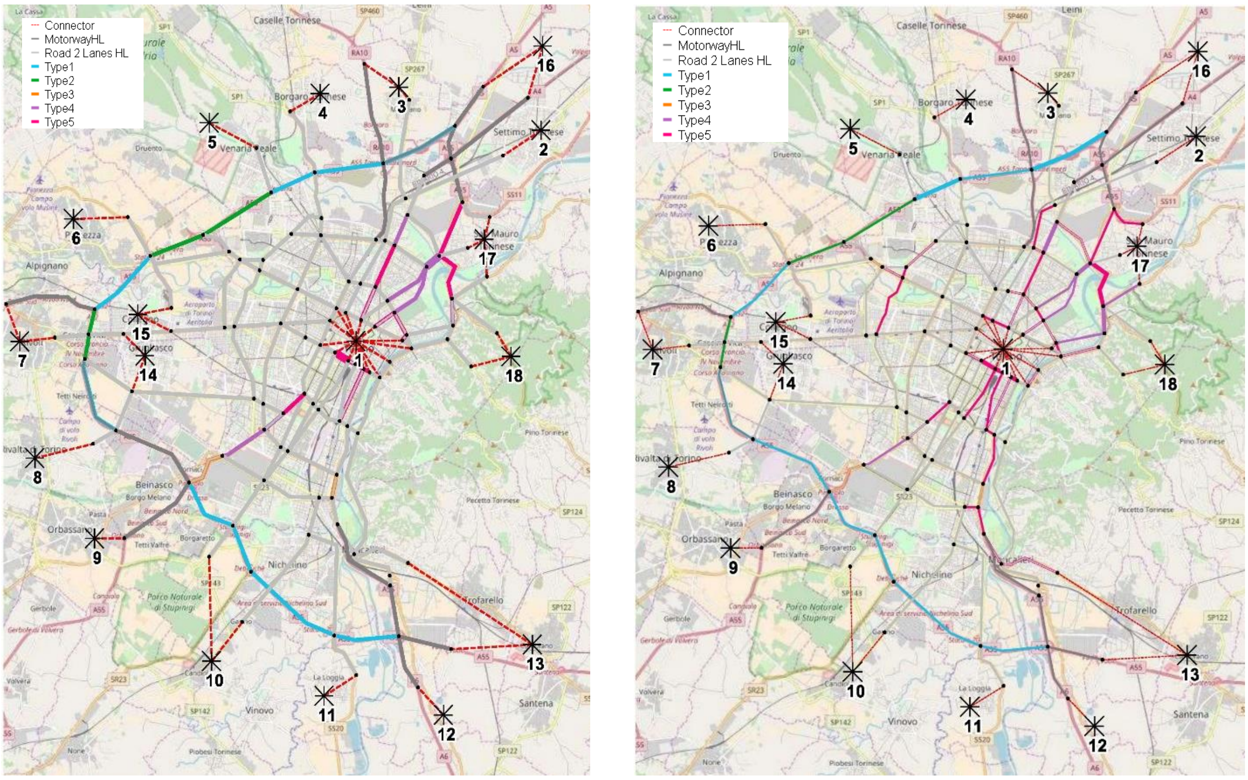
Table 1 Characteristics of the *a posteriori* network road type classification (time range H1).

New road type class	**Criteria [km/h]**	**Average [km/h]**	**No. arcs 10 values**	**No. arcs 5 values**
Type1	s* > 110	120	9	12
Type2	110 ≤ s < 80	105	4	5
Type3	80 ≤ s < 40	58	2	3
Type4	40 ≤ s < 20	29	9	13
Type5	s ≤ 20	10	14	44

* s: speed

In Fig. 8, the *a posteriori* network for H1 refines the one presented in Fig. 3 (*a priori*). More precisely, the figure contains the *a posteriori* network with its new classification for those arcs with at least 10 values of speed (Fig. 8(a)) and with at least five measures (Fig. 8(b)). A total of 38 updated (coloured) arcs are highlighted in the first case, while this number increases to 77 in the second. The comparison between the *a priori* and *a posteriori* network helps to increase knowledge of roads that are frequently travelled by the vehicles of this specific dataset. For example, it is important to note that many of the secondary arcs (those previously classified as “Road2lanes” in Fig. 3) are not as frequently covered by logistic fleet routes in Fig. 8(a), with the exception of those entering the city from the north-east and the south-west (violet and pink links). This meets expectations because the corresponding roads are along the connections between the areas around Turin where logistic structures are mainly located. Moreover, it is worth highlighting that the average travel speeds associated with those arcs are the lowest (29 or 10 km/h), as identified by the violet and pink coloured lines, representing somewhat congested streets. On the other hand, higher values are found for the Turin Ring road. In fact, both maps in Fig. 8 show cyan and green links for this road, meaning that the delivery vehicles travel at average speeds of 120 km/h and 105 km/h, respectively. These considerations are applied in the following evaluations of the results obtained by analysing the connections of pairs of centroids through shortest paths.

Although the dataset for H1 with more than 10 measures of speed has been used to classify the links of the *a posteriori* network, additional information could be gained considering a wider amount of links, including those with at least five values of speed, which account for 36% of the total links travelled by the fleet. Fig. 8(b) displays them on the map, providing a more detailed characterisation of the city centre compared to Fig. 8(a). This will be exploited in depth in the following sections to gain knowledge regarding city accessibility as perceived by the delivery fleet. A similar representation for the other time range (H2) is shown in Fig. 9. Here, the classification derived previously is applied and those arcs with at least five values of average speed are displayed and a total of 44 links is found. The matching of this map with the corresponding one for H1 (Fig. 8(b)) stresses that different roads are travelled in the two periods of the day by fleet vehicles. Moreover, a further variation is observed when comparing the average speed of some links composing the Turin Ring road. In fact, for H1 (Fig. 8(b)), higher speed values are detected (Type1 and Type 2, respectively 120 km/h and 105 km/h), whereas in H2 the average travel speed decreases to 58 km/h or even 10 km/h for some links (Fig. 9). During the late afternoon, these congested road conditions are familiar to frequent drivers, which is confirmed by the information extracted from the GPS traces dataset. In addition, these measures correspond to different days of the month, indicating that this situation is rather common and is not simply due to an unusual event, such as a car accident or the presence of road work. The choice of more than 10 speed measures should limit the influence of such random events in the estimated values.



(a) (b)

Fig. 8 *A posteriori* networks using arcs with at least (a) 10 measures and (b) 5 measures for the time interval H1 (Source: OmniTRANS)

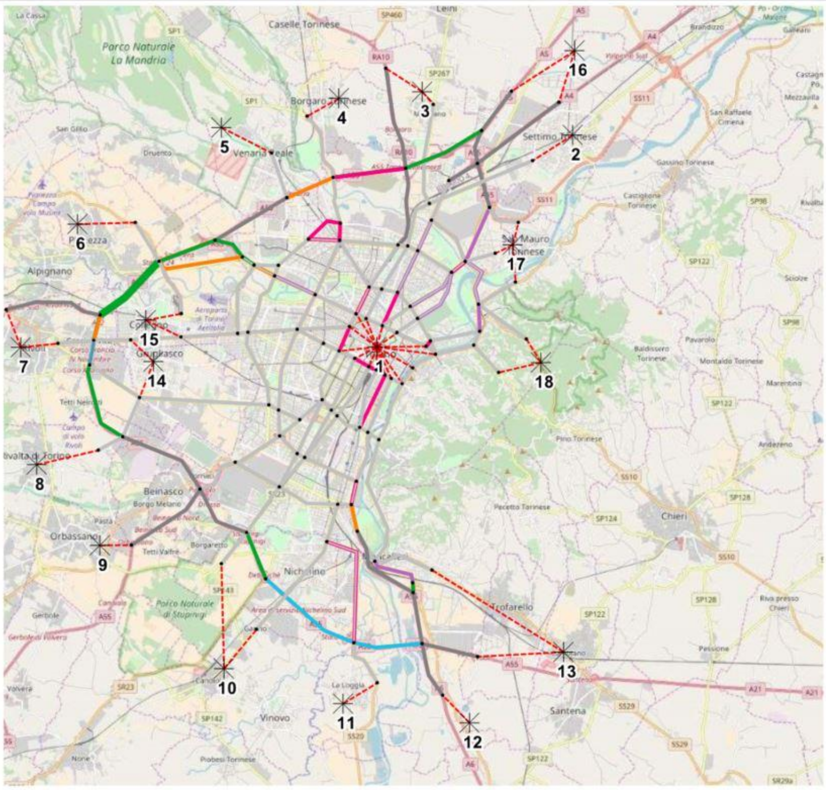


Fig. 9 *A posteriori* network using arcs with at least 5 measures for the time interval H2 (refer to Fig. 6 for the legends, source: OmniTRANS)

3.4. Verification and validation of a posteriori *network*

As explained in Section 3.3, the arcs of the network where data are available have been assigned to one of the five possible classes according to the computed average travel speed. As discussed in the Methodology sections, an alternative approach would require use of specific speed values estimated for each link to evaluate the accessibility to the zones in the study area. Hence, validation of the classification leading to the *a posteriori* network definition involves a comparison between these two scenarios through comparison of the time necessary to travel amongst the origin/destination (O/D) pairs of the network. In this test case, for all O/D pairs, the difference in travel time is less than 1 min, with the exception of some routes directed to zone 18 (less than 4 min), because of low speed links (Type 5) with higher travel times (Fig. 10). Therefore, the validity of the proposed classification is confirmed when approximating specific link values with respect to the accessibility estimation among zones.

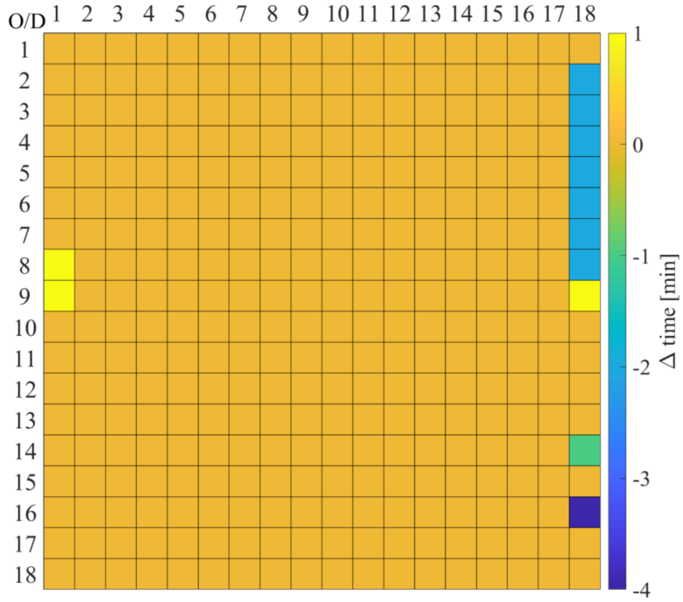


Fig. 10 Difference in the values of travel time [min] of shortest paths connecting 18 centroids using real speed versus the *a posteriori* H1 network (where speeds are represented by classes).

The length and travel time duration of the minimum paths for some selected connections are checked with respect to three applications available on the Web and operated by well-known map providers: Google maps (www.google.com/maps), Here data (www.here.com) and OpenStreetMap (www.openstreetmap.org). The travel time comparisons are presented in Table 2.

Table 2 Travel time [min] comparison for different routes between pairs of centroids using various commercial applications (see Fig. 11 for centroid positions)

Route	***A posteriori* network**	**Google Maps**	**Here**	**OSM**
10080-10066	22	18-28	25	20
10066-10080	28	20-35	28	21
10087-10106	36	24-50	33	31
10106-10087	27	24-50	30	32
10080-10114	34	26-45	36	33
10114-10080	25	26-50	34	32

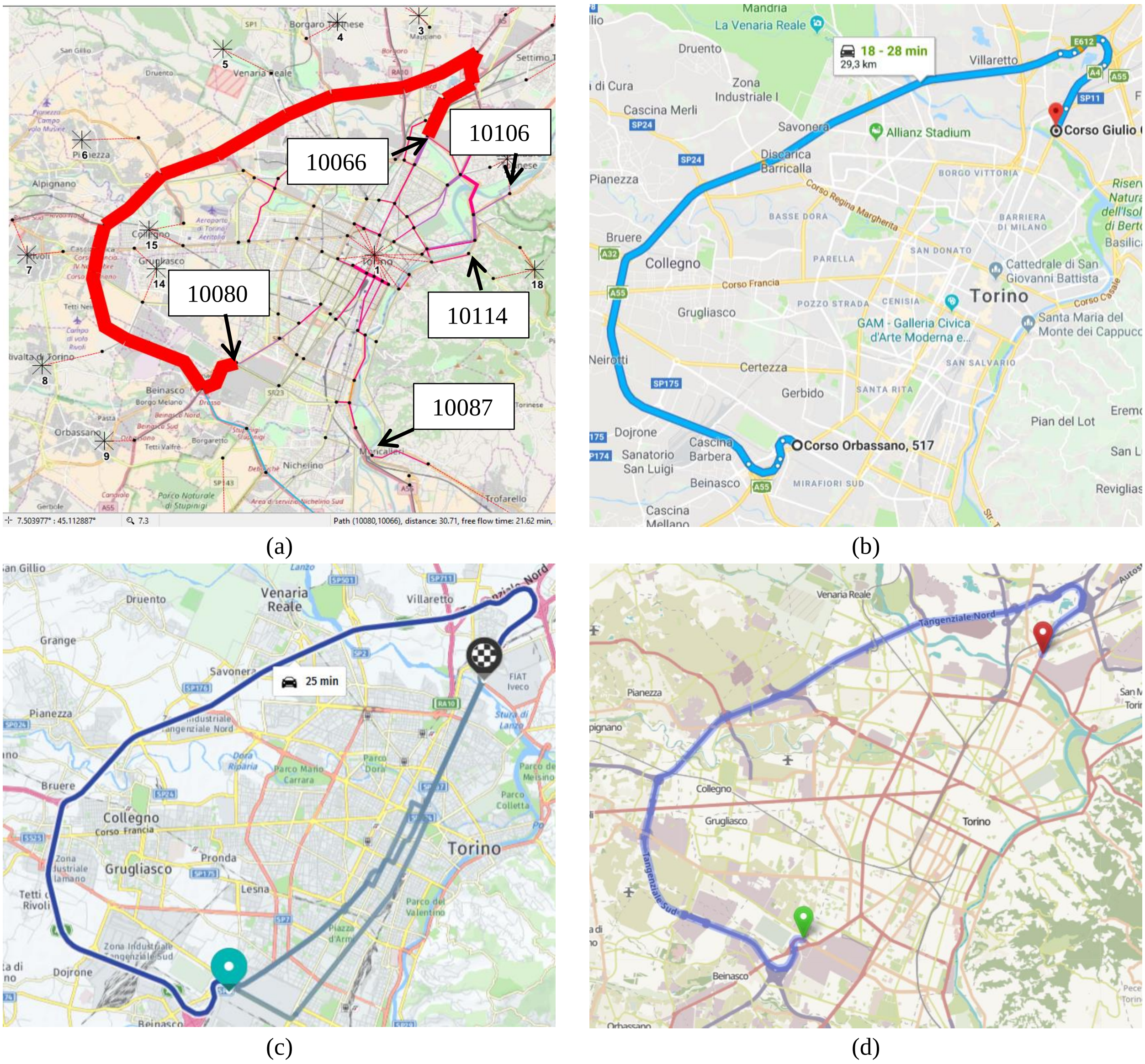


Fig. 11 Minimum path comparisons using different commercial applications (example of route from node 10080 to node 10066): (a) OmniTRANS, (b) Google Maps, (c) Here, (d) OpenStreetMap

As indicated in Table 2, each Web service provides different values for similar routes (Fig. 11). Google Map, for instance, specifies a range of travel time consistent with the one obtained with the presented methodology. One of the reasons behind these differences could lie in the vehicle types included in the travel time calculation. In our case, the recordings come from delivery vans, while other web services could also draw from other sources. It could be expected that their duties influence the speed of the former kind of vehicle, but, as presented in the methodology section, our procedure deals with removing the stop time required in those activities. The verification presented can indicate that the network model will not provide out of range values for travel times between relevant zone connections and the results that we obtain are consistent with those derived from other tools.

4. Discussion on the accessibility results

4.1. Time and distance influence on route selection

The relevance of time for the best route selection can be measured by comparing the travel length for the shortest paths found (considering time as the link attribute) to the length of the path between the same pair of centroids on the basis of distance attributes. Fig. 12 visualises these differences between lengths obtained for the two types of path calculations, considering each time range (H1 and H2) separately. It is interesting to note the polarisation of greater variations in certain zones, meaning that the contribution of the GPS dataset has a relevant influence on the travel time necessary to go from specific centroids to others. However, the absence of a complete refinement of the network has a definite impact, as no information could be added to a more "static" component of the network, such as the distance, which is computed based on the lengths of the arcs. For example, assuming time as the attribute, although the length of the path from 13 to 17 is 23 km greater than the case of assuming distance, 18 min have been saved, as shown in the first row of Table 3. Fig. 13 shows the changes in this path, presumably as a result of congestion and the refined information regarding the travel speed contained in the *a posteriori* network. It is interesting to note that in the second time range (H2), the major change is symmetrical to the case of H1 (last two rows of Table 3). The information derived by such type of value analysis could provide useful insight as to the level of efficiency of the network. In fact, if the reduction of time necessary to connect two centroids is associated with an increase in the kilometres travelled, this would imply greater consumption of resources by the vehicles related to the distance, such as fuel or tyres.

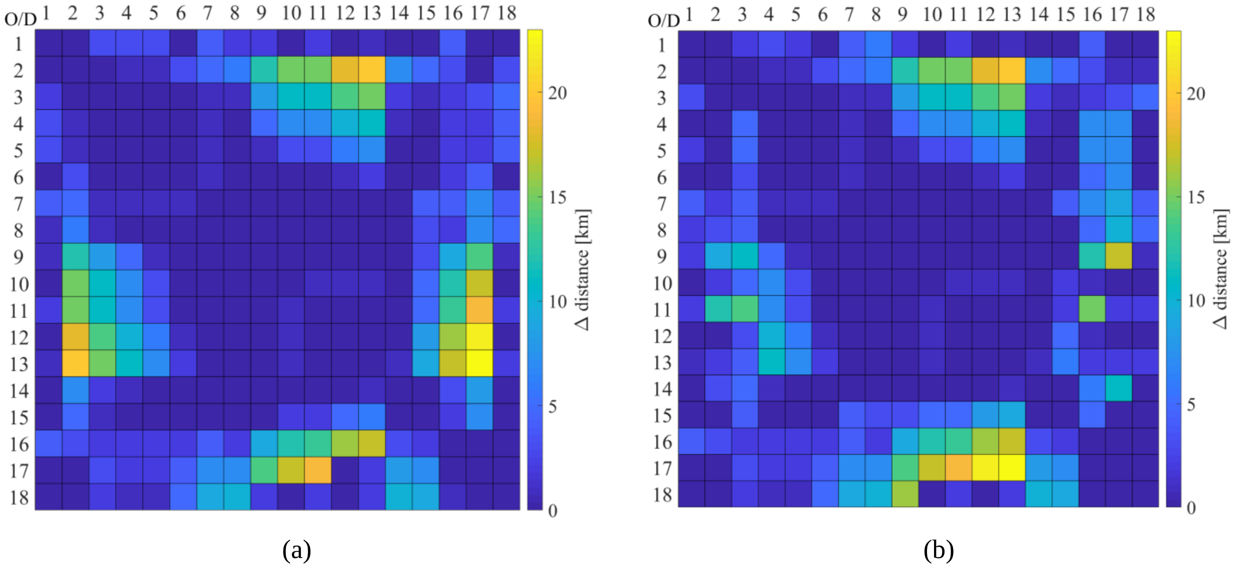
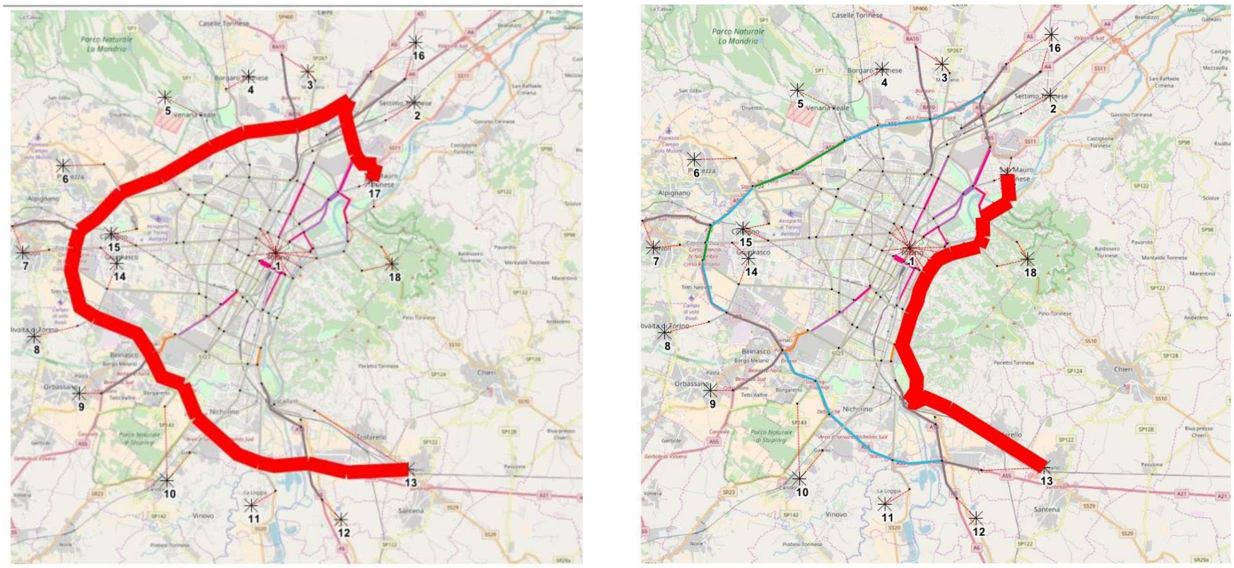


Fig. 12 Length variation [km] of shortest path for time/distance-based algorithm in the two time ranges: (a) H1 and (b) H2

Table 3 Shortest paths comparison for a particular route (centroids 13-17)

	From	**To**	**Distance_T* [km]**	**Travel time_T* [min]**	**Distance_D* [km]**	**Travel time_D* [min]**	**Δ distance [km]**
H1	13	17	46.3	42	23.0	60	23.3
	17	13	24.7	49	23.0	60	1.7
H2	13	17	24.7	50	23.0	60	1.7
	17	13	46.3	48	23.0	60	23.3

*"T" = shortest path based on the travel time / "D" = shortest path based on the distance



(a) (b)

Fig. 13 Different shortest paths connecting centroids 13 and 17 for time range H1: (a) 46.3 km and (b) 24.7 km (source: OmniTRANS).

4.2. Skim matrices comparison for the two time periods

The influence of floating car data (FCD) integration on the travel time matrices is highlighted in this section by considering that in different time periods, the speed may change on the congested links. In Fig. 14, the differences of travel time between the best paths (selected on the basis of the time attributes) of the two time periods are depicted for the various zones. The highest value corresponds to approximately 15 min and the same connections can be slower or faster for the two periods, depending on the pairs of zones.

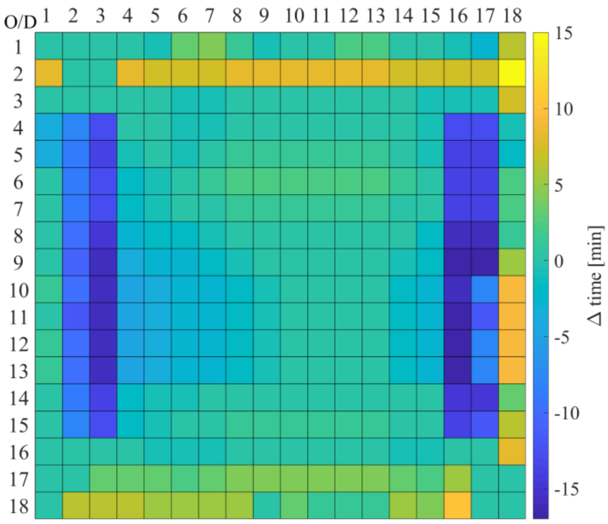
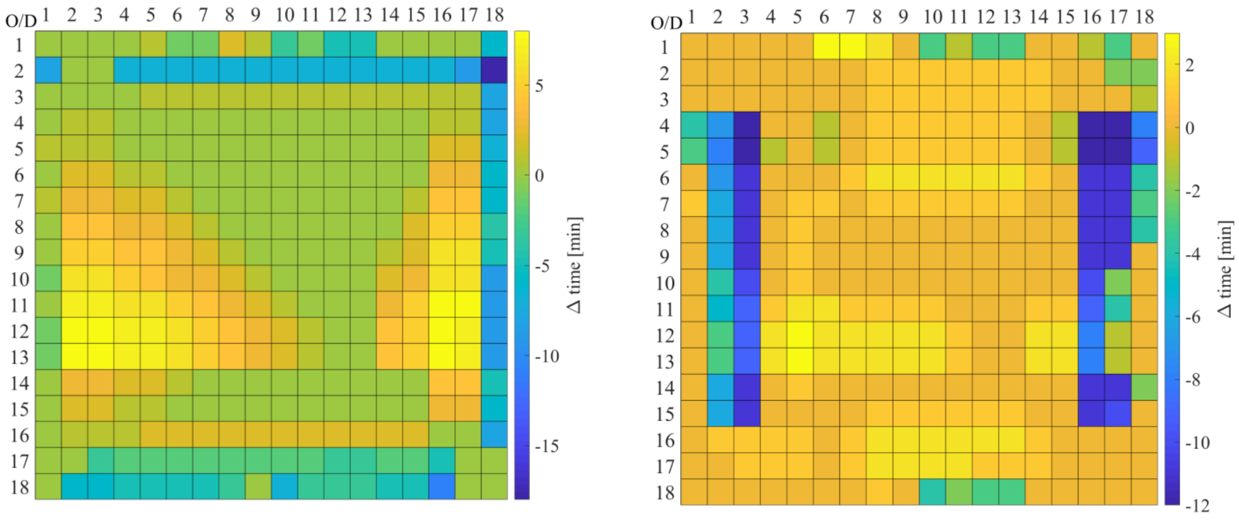


Fig. 14 Travel time difference [min] between time ranges H1 and H2 for the scenarios of the *a posteriori* network with at least 5 measures.



(a) (b)

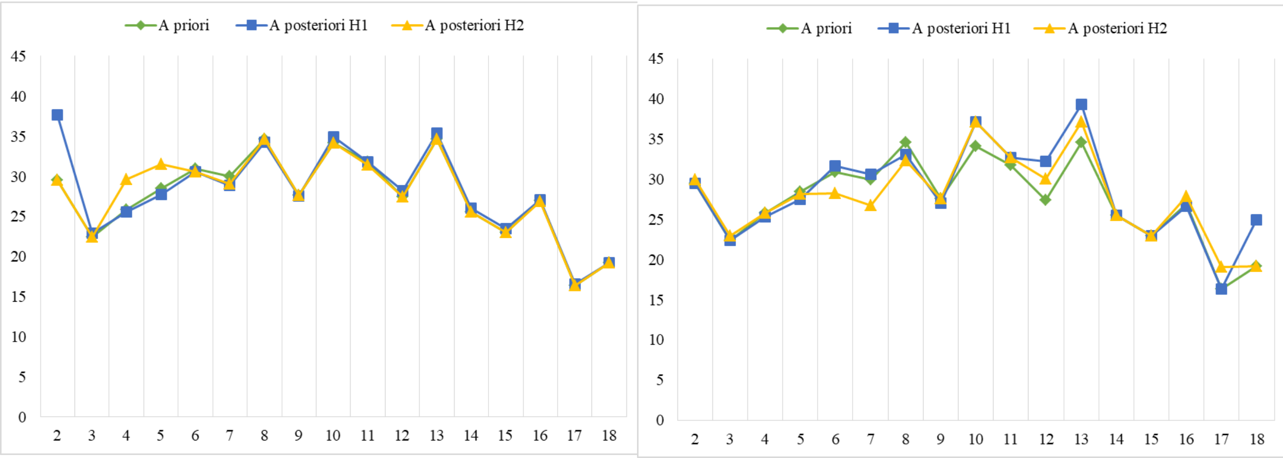
Fig. 15 Travel time difference [min] between the *a priori* and *a posteriori* scenarios for time range: (a) H1 and (b) H2. In both cases, the *a posteriori* network with at least 5 measurements is considered.

The valued added by the refinement process on the network model by separating the time periods is confirmed in Fig. 15, where the differences are apparent by comparing the travel time as estimated by the skim matrix for the *a priori* and *a posteriori* networks of H1 and H2. In both cases, a low value, seen as a difference and not as an absolute value, corresponds to more time required to connect a specific pair of centroids in the *a posteriori* scenario. For example, as confirmation of the discussion in Section 3.3, paths reaching centroid 18 in Fig. 15(a) are usually associated with negative values mainly because most of the nearby links are characterised by low speed values in the *a posteriori* case, as can be observed by comparing Fig. 3 and Fig. 8(a). However, the richness given by the knowledge derived with the refinement of the *a priori* network is confirmed by the fact that 77% of the values are different from zero in both cases.

4.3. Insight on accessibility for specific zones

A further challenging application of the method focusses on the measurement of the accessibility to and from crucial centroids for delivery operations, such as the city centre for its business relevance, the connections with external metropolitan areas, or the zones where depots are located. This information, in terms of travel time, may be helpful to properly plan delivery trips by fleet managers or to support location decisions for logistic structures within a city. In fact, the knowledge of the shortest paths for different network configurations (in H1 and H2 time periods in this case) could provide interesting feedback on the accessibility of various zones under investigation.

A first focus could be the city centre, i.e. centroid 1, as origin (Fig. 16(a)) or destination (Fig. 16(b)) of all possible connections with other centroids. For instance, in the first case, few variations in values are found, meaning that both the time range and the refinement of the *a priori* network seem to have minor influence on the travel time when the routes are oriented towards the city centre.

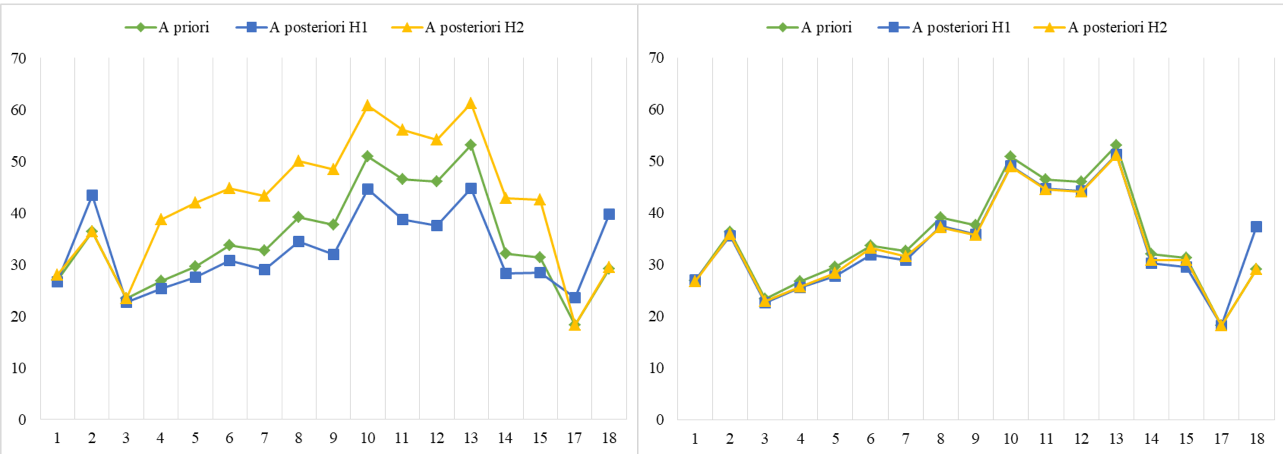


(a) (b)

Fig. 16 Comparison between the travel time [min] for three different scenarios involving: (a) to and (b) from the city centre (centroid 1).

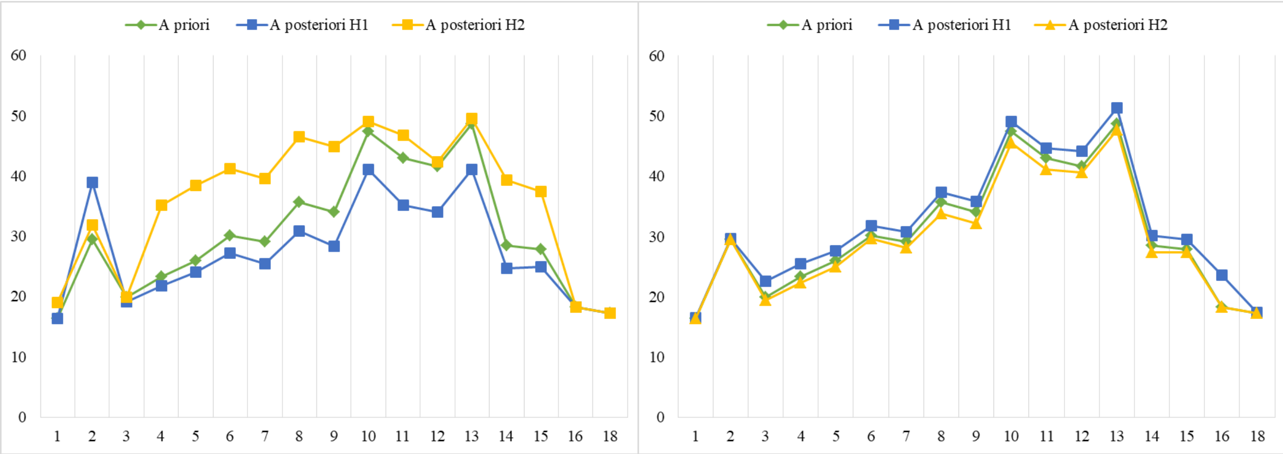
The accessibility of the Motorways located in the northern part of the city (i.e. to and from Milan) is another example of applying the proposed method to measure the quality of the network with respect to external stakeholders, as shown in Fig. 17. In this case, reaching other zones is strongly influenced by the time period, as confirmed by the differences between H1 and H2, as well as by the refinement process of the network with respect to the *a priori* design. Finally, similar charts are shown in Fig. 18 for centroid 17, which approximates the position of the area where a cluster of depots managed by freight distribution companies is currently located. Most of the variations are found, as for the previous centroid, for travel along paths connecting to that specific zone rather than for those leaving it, as shown in Fig. 18(a) and (b), respectively.

It is worth observing that the assumptions adopted for the speed values in the *a priori* network produce travel time values in Fig. 17 and Fig. 18 that are intermediate between those in H1 and H2. This reveals that the authors' knowledge of the average speed used to preliminarily classify the links seems to be affected by the average traffic conditions in the two periods.



(a) (b)

Fig. 17 Comparison between the travel time in three different scenarios involving: (a) to and (b) from Motorway North (centroid 16)



(a) (b)

Fig. 18 Comparison between the travel time in three different scenarios involving: (a) to and (b) from the depots area (centroid 17)

Conclusion

The methodology presented in this paper was developed to investigate accessibility in urban areas, as perceived by freight distribution services operating using delivery vans, through the estimation of the average speed in a road network model. The low time resolution (sampling rate) of positioning data affected the model design, because not all road links connecting consecutive intersections could be included, only relevant arterials. For this reason, an approach based on nodes was adopted to detect vehicle positions along their travel routes, whereas the link information was used only to check a vehicle's direction. A classification of the links was also performed to simplify the model management based on the estimated speed between nodes of the road network.

The main results deriving from the case study in Turin confirm that the FCD values available from common commercial services can be used to improve the knowledge of the road network performance for various applications. In this study, high-level accessibility matrices were built to compare different zones of the city interested in delivery operations by analysing the most used urban connections. These first results are related to a specific type of trajectory data, collected by light duty vehicles, and they depict a well-defined situation that could be undoubtedly informative for certain stakeholders, such as public authorities and urban logistics operators.

The comparison of the travel time connecting different areas is another important characteristic to take into account. From the results of the application, for example, Motorway North (centroid 16) can be reached from some zones in approximately 20 min in the morning, but this value increases to 35 min if the same route is taken in the afternoon. However, the period does not significantly affect accessibility when considering travel in the opposite direction. This kind of information could provide useful suggestions on the creation of specific time ranges that could be exploited efficiently for the delivery operations along the day.

The proposed method and the derived accessibility matrices can be exploited by Local Authorities to obtain a global picture of the current network performance for management purposes. Furthermore, better knowledge of different scenarios can support the planning of future measures to regulate urban freight deliveries. The monitoring of the accessibility can help in the validation of reversible measures proposed at the city level such as: (a) use of reserved lanes also for delivery vans, (b) regulation of the time ranges for entering the city centre, (c) exploitation of special permissions.

Other stakeholder categories that might benefit from the results, such as travel time to reach an established zone, are van/fleet operators. They could exploit such findings as support for: (a) their decisions when planning delivery routes, (b) choosing the optimal time range(s) for parcel distribution by shifting from congested to off-peak periods, and (c) providing more accurate delivery time windows to end users. Overall, Local Authorities have to be able to access and manage this kind of information because they are expected to take into account the needs of different stakeholders acting in the field to be sure of creating the proper strategy for freight transport at the city level.

Based on the proposed framework, future work could try to apply the methodology by extending the focus to other urban areas where deliveries or city logistic operations are relevant. Besides, the availability of a more extensive database and integrating the trajectories of more freight operators could extend the knowledge pertaining to urban accessibility. Targeted analyses could also focus on different period of the year (summer/winter) or days of the week, to identify particular trends. The power of the approach proposed lies in the possibility of evaluating and monitoring the effects of reversible actions proposed at the city level (access in certain areas, use of reserved lanes, etc.) that would require a simulation model not always easily to be implemented. With all these aims, a network modelling tool, although here applied with only a small portion of its functionalities, could be used to manage additional associated information, such as traffic flow on links or the travel demand between specific zones.

References

ALICE, & ERTRAC. (2015). *Urban Freight research roadmap*. http://www.ertrac.org/uploads/documentsearch/id36/ERTRAC_Alice_Urban_Freight.pdf

Ambrosini, C., Patier, D., & Routhier, J. L. (2010). Urban freight establishment and tour based surveys for policy oriented modelling. *Procedia - Social and Behavioral Sciences*, *2*(3), 6013–6026. https://doi.org/10.1016/j.sbspro.2010.04.015

Ambrosino, G., Pettinelli, I., Freitas, C., & Sousa, C. (2015). Sustainable Urban Logistics Plan (SULP) methodology for Small and Mid-sized European Towns: the IEE ENCLOSE project results. *URBE 2015 URban Freight and BEhavior Change*, 1–12. http://www.polisnetwork.eu/uploads/Modules/PublicDocuments/1d_ambrosino_enclose_project.pdf

Ben-Akiva, M. E., Toledo, T., Santos, J., Cox, N., Zhao, F., Lee, Y. J., & Marzano, V. (2016). Freight data collection using GPS and web-based surveys: Insights from US truck drivers' survey and perspectives for urban freight. *Case Studies on Transport Policy*, *4*(1), 38–44. https://doi.org/10.1016/j.cstp.2015.11.005

CIVITAS WIKI consortium. (2015). *Smart choices for cities - Making urban freight logistics more sustainable*. Civitas Policy Note. http://www.eltis.org/resources/tools/civitas-policy-note-making-urban-freight-logistics-more-sustainable

Cui, J. X., Liu, F., Janssens, D., An, S., Wets, G., & Cools, M. (2016). Detecting urban road network accessibility problems using taxi GPS data. *Journal of Transport Geography*, *51*, 147–157. https://doi.org/10.1016/j.jtrangeo.2015.12.007

Curl, A., Nelson, J. D., & Anable, J. (2011). Does accessibility planning address what matters? A review of current practice and practitioner perspectives. *Research in Transportation Business and Management*, *2*, 3–11. https://doi.org/10.1016/j.rtbm.2011.07.001

de Palma, A., & Lindsey, R. (2011). Traffic congestion pricing methodologies and technologies. *Transportation Research Part C: Emerging Technologies*, *19*(6), 1377–1399. https://doi.org/10.1016/j.trc.2011.02.010

Diana, M., Pirra, M., & Woodcock, A. (2020). Freight distribution in urban areas: a method to select the most important loading and unloading areas and a survey tool to investigate related demand patterns. *European Transport Research Review*. https://doi.org/10.1186/s12544-020-00430-w

ELTISplus. (2017). *Guidelines. Developing and Implementing a Sustainable Urban Mobility Plan*.

Fossheim, K., & Andersen, J. (2017). Plan for sustainable urban logistics – comparing between Scandinavian and UK practices. *European Transport Research Review*, *9*(52), 1–13. https://doi.org/10.1007/s12544-017-0270-8

Freight Leaders Council. (2017). *Quaderno 26 - La logistica ai tempi dell'e-commerce*.

Fu, J., & Jenelius, E. (2017). TRANSPORT EFFICIENCY OF OFF-PEAK URBAN GOODS DELIVERIES: A STOCKHOLM PILOT STUDY. *Journal of the Transportation Research Board*, 1–15.

Ge, Q., & Fukuda, D. (2016). Updating origin-destination matrices with aggregated data of GPS traces. *Transportation Research Part C: Emerging Technologies*, *69*, 291–312. https://doi.org/10.1016/j.trc.2016.06.002

Geurs, K. T., & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: Review and research directions. *Journal of Transport Geography*, *12*(2), 127–140. https://doi.org/10.1016/j.jtrangeo.2003.10.005

Gonzalez-feliu, J., Pluvinet, P., Serouge, M., & Gardrat, M. (2013). GPS-based data production in urban freight distribution. *Global Positioning Systems: Signal Structure, Applications and Sources of Error and Biases*, 1–20.

Greaves, S. P., & Figliozzi, M. A. (2008). Commercial vehicle tour data collection using passive GPS technology: issues and potential applications. *TRB 87th Annual Meeting Compendium of Papers, Washington D.C.*, 1–18.

Hess, S., Quddus, M., Rieser-Schüssler, N., & Daly, A. (2015). Developing advanced route choice models for heavy goods vehicles using GPS data. *Transportation Research Part E: Logistics and Transportation Review*, *77*, 29–44. https://doi.org/10.1016/j.tre.2015.01.010

Holt, D. H., Author, C., Beach, L., & Beach, L. (2017). Analyzing Truck Traffic in Mississippi via GPS Transponders. *Presented at 96th Annual Meeting of the Transportation Research Board, 8-12 January 2017, Washington D.C., USA*.

Kiba-Janiak, M. (2017). Urban freight transport in city strategic planning. *Research in Transportation Business and Management*. https://doi.org/10.1016/j.rtbm.2017.05.003

Mcnally, M. G. (2007). The Four Step Model. In D. . Hensher & K. . Button (Eds.), *Handbook of Transport Modeling* (pp. 35–52). Pergamon, Oxford. https://doi.org/https://escholarship.org/uc/item/0r75311t

Meyer, A., & Meyer, D. (2013). City Logistics Research: A Transatlantic Perspective. In *Transport Research Symposium*. National Academy of Sciences. https://doi.org/10.17226/22456

Pascale, A., Deflorio, F., Nicoli, M., Dalla Chiara, B., & Pedroli, M. (2015). Motorway speed pattern identification from floating vehicle data for freight applications. *Transportation Research Part C: Emerging Technologies*, *51*(305), 104–119. https://doi.org/10.1016/j.trc.2014.09.018

Patire, A. D., Wright, M., Prodhomme, B., & Bayen, A. M. (2015). How much GPS data do we need? *Transportation Research Part C: Emerging Technologies*, *58*, 325–342. https://doi.org/10.1016/j.trc.2015.02.011

Pirra, M., & Diana, M. (2019). Integrating mobility data sources to define and quantify a vehicle-level congestion indicator: an application for the city of Turin. *European Transport Research Review*. https://doi.org/10.1186/s12544-019-0378-0

Pluvinet, P., Gonzalez-Feliu, J., & Ambrosini, C. (2012). GPS Data Analysis for Understanding Urban Goods Movement. *Procedia - Social and Behavioral Sciences*, *39*, 450–462. https://doi.org/10.1016/j.sbspro.2012.03.121

Pronello, C., Camusso, C., & Valentina, R. (2017). Last mile freight distribution and transport operators' needs: which targets and challenges? *Transportation Research Procedia*, *25*, 888–899. https://doi.org/10.1016/J.TRPRO.2017.05.464

Sharman, B. W., & Roorda, M. J. (2013). Multilevel modelling of commercial vehicle inter-arrival duration using GPS data. *Transportation Research Part E: Logistics and Transportation Review*, *56*, 94–107. https://doi.org/10.1016/j.tre.2013.06.002

Taylor, M. A. P., Woolley, J. E., & Zito, R. (2000). Integration of the global positioning system and geographical information systems for traffic congestion studies. *Transportation Research Part C: Emerging Technologies*, *8*(1–6), 257–285. https://doi.org/10.1016/S0968-090X(00)00015-2

Yang, X., Sun, Z., Ban, X., & Holguín-Veras, J. (2014). Urban Freight Delivery Stop Identification with GPS Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2411*, 55–61. https://doi.org/10.3141/2411-07