Taylor & Francis
Taylor & Francis Group

Journal of Maps

ISSN: (Print) 1744-5647 (Online) Journal homepage: https://www.tandfonline.com/loi/tjom20

From infrastructure to service: mapping long-distance passenger transport in Italy

Paolo Beria, Andrea Debernardi, Raffaele Grimaldi, Emanuele Ferrara, Antonio Laurino & Alberto Bertolin

To cite this article: Paolo Beria, Andrea Debernardi, Raffaele Grimaldi, Emanuele Ferrara, Antonio Laurino & Alberto Bertolin (2016) From infrastructure to service: mapping long-distance passenger transport in Italy, Journal of Maps, 12:4, 659-667, DOI: 10.1080/17445647.2015.1060179

To link to this article: https://doi.org/10.1080/17445647.2015.1060179

JOURNAL OF MAPS, 2016
VOL. 12, NO. 4, 659–667
http://dx.doi.org/10.1080/17445647.2015.1060179

SOCIAL SCIENCE

From infrastructure to service: mapping long-distance passenger transport in Italy†

Paolo Beria[a], Andrea Debernardi[b], Raffaele Grimaldi[a], Emanuele Ferrara[b], Antonio Laurino[a] and Alberto Bertolin[a]

[a]Dipartimento di Architettura e Studi Urbani, Politecnico di Milano, Via Bonardi 3, 20133 Milan, Italy; [b]Studio META, Via Magenta 15, 20090 Monza, Italy

ABSTRACT
In order to promote an effective level of coordination between physical investments, technology and soft policies in transport planning, a deep knowledge of supply and demand is desirable, if not necessary. Unlike other countries, the national scale of supply and demand for the Italian transport systems as a whole is barely known and in the case of long-distance mobility, there is not a unique quantitative and geographical description available. In this paper, we present a map regarding the Italian long-distance transport supply and generalised cost simulations, for the period 2013–2014. The information shown in the map comes from a multimodal transport model, which presents the peculiarity of using real public service timetables to simulate the entirety of the Italian long-distance transport industry. This tool enables one to map the entire transport supply and to estimate the generalised costs among any route: this also allows one to identify which transport mode is better suited to make a specific trip.

ARTICLE HISTORY
Received 7 October 2014
Revised 6 May 2015
Accepted 26 May 2015

KEYWORDS
Long-distance transport industry; transport model; generalised cost

1. Introduction

Transport planning at any scale is commonly based, in many countries, on a deep knowledge of the current demand and on a detailed description of all dimensions of existing infrastructure, the supply of services (for public transport) and of market conditions.

However, in Italy, the national scale of transport is barely known from a quantitative and geographical point of view. In the case of long-distance mobility, a complete picture is not known at all, except for specific components such as commuters mobility (ISTAT, 2014) or air routes flows (Assaeroporti, 2014; ENAC, 2014) or the aggregated data of the 'Conto Nazionale delle Infrastrutture e dei Trasporti' (Ministero delle Infrastrutture e dei Trasporti, 2014). This single-mode vision is furthermore usually based on infrastructure investments and ignoring the transport services, and misses the opportunity of a more rich coordination of physical investments, technology and soft policies, such as pricing (Boitani & Ponti, 2006).

In order to provide a useful tool to ride this lack of knowledge out, Studio META and the Research Centre on Transport Policy (TRASPOL) collaborated to develop a complete quantitative and spatially defined description of all long-distance Italian passenger transport, including both infrastructure and services.

In this section, we outline the different components and the operative procedures used to collect transport supply data (rail, long-distance bus and air services and related timetables) used to feed the georeferenced multimodal transport simulation model of the entire Italian transport industry in 2013–2014. The database has been used to prepare some charts of Italian transport supply characteristics, which are the core of this paper and which will be described extensively in the following parts.

Extensive reviews of different techniques and tools used in transport modelling can be found in Cascetta (2006). The peculiarity of this model is that it has both a strong and detailed spatial dimension (sub-provincial zoning, georeferenced infrastructures, etc.), and in order to simulate real interchanges among different modes, it includes the 2013–2014 real service timetables.

Two modules of calculation determine the main structure of the model. The first module (supply module) allows one to estimate the matrices of travel time, operating costs and fares of transportation for every origin/destination (O/D) relation and for each modal option taken into account. The second module (modal split module) allows estimating the generalised cost of travel and, after calibration, the consequent probable modal choice for different users' profiles.

†The paper is the outcome of a work involving all authors, focused on the overall developing of the Italian National Transport Model. Despite the collaborative structure and the common responsibility, tasks were divided as follows. Andrea Debernardi and Emanuele Ferrara mainly developed the model (database structure, mathematical algorithms, etc.) and the infrastructure database. Paolo Beria, Raffaele Grimaldi and Antonio Laurino were mainly responsible for services database (rail, coach, air and fares). Alberto Bertolin produced the Main Map and, together with Paolo Beria, wrote the paper. The model did not receive any direct financing and was autonomously developed by the authors.

CONTACT Paolo Beria paolo.beria@polimi.it

In transport economics, the generalised cost is the sum of monetary and non-monetary costs perceived by the user to perform a certain trip. This measure represents

> the minimum cost of transporting a given load of a particular commodity between a specific origin and a destination, considering the economic variables related to the input costs necessary to produce the transportation service, and the physical features of the available transport infrastructure. (Zofío, Condeço-Melhorado, Maroto-Sánchez, & Gutiérrez, 2014, p. 142)

Generalised cost is used, in the calibrated modal split and assignation modules (not covered by this paper), to determine the modal shares and the chosen paths by means of discrete choice models (Cascetta, 2006; Ortúzar & Willumsen, 1990).

2. Methods

2.1. *Preparation of cartographical databases*

2.1.1. *Sub-provincial zoning*

The adopted sub-provincial zoning includes 371 zones. Each zone identifies a traffic catchment area that generally represents a homogeneous aggregation of municipalities on the base of their population. Each aggregation is also comparable with the European statistical level NUTS 4 (European Commission, 2007) which does not have a direct correspondence with any Italian administrative boundary. Table 1 and Figure 1 show the total number of zones in which each Italian region was divided.

With this structure, the model is able to simulate 371 × 371 = 137,641 O/D relations. This level of detail is necessary to describe both the complexity of Italian transport system and the population distribution and interrelations. Instead, the official European statistical level with highest disaggregation (NUTS 3, which subdivides Italy in 107 zones) would not have been sufficient to describe the mobility of areas as large as 2–3000 km^2 on average.

Table 1. Number of zones per region.

Regions	N. provinces	N. zones
Piemonte	8	29
Valle d'Aosta	1	1
Lombardia	12	46
Trentino-Alto Adige	2	10
Veneto	7	25
Friuli-Venezia Giulia	4	11
Liguria	4	8
Emilia-Romagna	8	28
Toscana	10	31
Umbria	2	9
Marche	6	15
Lazio	5	27
Abruzzo	4	10
Molise	2	3
Campania	5	21
Puglia	6	28
Basilicata	2	7
Calabria	5	12
Sicilia	9	32
Sardegna	8	18
Total	110	371

2.1.2. *Infrastructure*

The infrastructural network database is derived from the aggregation of different sections of official regional cartographies CTR (Carta Tecnica Regionale). These types of topographic maps (available both as raster and vector file) are produced independently by the Italian regions in order to represent (in a scale between 1:5000 and 1:10,000) their administrative territory and to show the real projection and shape of each punctual elements. Every region prepared different territorial recognitions, also running several years between one and the other, to elaborate its own cartography. Consequently, in order to obtain up to date and homogeneous information, it was necessary firstly to merge CTRs, with integrations from updated web sources. From this cartography, it was finally possible to reconstruct in detail the main road, rail, naval and airport connections. Nodes and links are at the base of each sub-element of the graph and connect them with each other and with the 371 traffic zones.

In detail, the multimodal graph includes five principal network classes. The first class represents the entire national *railway network* composed by 4052 nodes and 4500 bidirectional links. The national railway shapefile is enriched by further information based on official data of operators: number of tracks, gauge, control system, station description, number of intersections, etc. (Figure 2).

The second class identifies the entire national *road network* (subdivided in highway, provincial road and main connections at the sub-provincial level). The shapefile includes a double classification for links. The first description is based on the geometric characteristics of the road (number of lanes, intersections, etc.), while the second one refers to the role of the road as a connection (i.e. connecting NUTS 2 zones, etc.) and is used for modelling purposes. In order to subsequently calculate the generalised cost of private mode of transport, each edge includes also information on the average speed (from 120 to 20 km/h) allowed on the base of endogenous (type of road) and exogenous (orography and urban contest) elements, determined in part automatically and in part manually (Figure 3).

The third class is *the maritime and internal navigation network*. This shapefile includes only routes used by ferry services. These links are used to ensure the continuity of the road network between the mainland and the islands. We built both 'navigation' and 'boarding' links: the first ones are used to simulate the average speed of the ship, while the second ones identify the time consumption for boarding/landing operations.

The fourth class is the representation of the *air navigation network*. The air navigation shapefile includes three different types of links, which overcome the

Figure 1. Zoning adopted for the model.

pure geographical description and are used for modelling purposes. Similar to the maritime shapefile, the first link corresponds to the connection between two airports, the second one identifies both time consumption and disutility for land–air interchange (parking fee, time needed to find a parking space, etc.) and the third refers to time for check-in/check-out operations. This third type of link also presents a sub-classification linked to the type of flight chosen by different user categories (low-cost flights or full cost flights). Due to this complex structure, this shapefile includes two types of nodes; the first ones represents the airport as a physical element accessible by the road or railway networks, while the second one is a virtual representation of gates and it is used to connect check-in/check-out borders with those of navigation (Figure 4).

The last class includes all the *zonal and intermodal connectors* that provide the link between transport demand data (organised in an Access database) and the infrastructure graph. Thanks to this architecture, the transport model described in this paper is more realistic to determine the real generalised costs of trips with private vehicles (see Section 2.3.1).

In order to avoid miscalculations in computing the minimum paths for each O/D relation, a semi-automatic debug procedure was implemented to identify and subsequently correct possible connectivity errors (e.g. lack of continuity between consecutive sections of the same road or railway).

2.2. *Preparation of timetable databases*

The second step of model database definition requires the implementation of a timetable database and its connection with infrastructure networks. This effort makes the model able to describe also the generalised costs associated with the existing public transport supply operating on medium and long-distance O/D relations. In the following paragraphs, we define the sources that have allowed us to set up the timetables of public

Figure 2. Italian rail network. In black, the national rail network defined as 'fundamental' and, in blue, the rest of the lines, including those managed by local operators different from the national one. Closed lines are included, too.

transport services and, subsequently, the methodology used to visualise this information into a georeferenced hypergraph.

2.2.1. *Coach service timetables*

The timetable database contains a complete description of Italian long-distance coach services (average winter weeks of 2013–2014). All information on routes, stops and frequency of trips was derived directly from the websites of the numerous coach companies. In total, 391 long-distance bus lines have been modelled, operated by 80 different operators. This information, related to an average weekday, was subsequently standardised into database environment to allow comparison and data elaboration. In order to allow the model to simulate interchanges, two columns were added in the database, namely a binomial column related to the presence of stops in common with more transport system (for instance, bus stops close to airports) and another with the day of the week in which the service is provided.

2.2.2. *Rail service timetables*

The timetable database contains a description of any Italian rail services (average winter week of 2014). At the current stage of implementation, for the regional rail services, the database includes only those routes provided by the primary national operator Trenitalia, and some local rail companies (Tper, Trenord, etc.). Some local concessions are excluded, but their degree of integration with long-distance transport is extremely limited. Instead, the totality of long-distance services is included, and provided by the two existing operators, Trenitalia and NTV, for 560 trains per day (2014). All information on routes, stops and frequency of the trips was derived directly from the websites and official timetables of transport companies. For this database,

Figure 3. National road network.

we utilised the same procedure and standardisation scheme adopted for the coach services.

2.2.3. *Air service timetables*

The timetable database contains a complete description of domestic Italian air routes (average spring week of 2013), including 280 single routes. All information on routes, stops and frequency of trips was derived directly from the OAG Database. Also in this case, we utilised the same procedure and standardisation scheme adopted for land transport.

2.3. *Generalised cost calculation*

According to Nichols (1975), generalised costs measure depends firstly on distance and time and it represents a translation of this key accessibility variables in economic costs (units prices).

As different individuals perceive differently the cost components of a trip, we defined three stylised demand segments: the business travellers (which tend to prefer fast modes, such as the car or the plane), the economy travellers (which do not have a car and tend to reduce monetary costs in change of longer travel time) and the families (which are 'economy' travellers, but can share the cost of a private car). Each segment is defined in the simulations by four characteristics, as summarised in Table 2.

2.3.1. *Calculation of the generalised cost for private vehicles*

Concerning the private road transport, the generalised cost utilised in this model derives from the usual definitions (Ortúzar & Willumsen, 1990) and it is calculated for a single road edge using the following formula:

$$G_C = aD + bTN + cP,$$

where D is the distance (km). T is the time required to travel that distance (km/h) on the base of the average speed allowed on that specific arc. P is the toll, where applicable (typically on motorways). In the case of a

Figure 4. Air navigation network.

closed tariff system (as in the majority of Italian motorways), it is calculated as the sum of a kilometric toll (0.05€/km) plus a fixed rate (0.50€), while, in the few cases of an open tariff system, it is equal to the current real toll applied at the gates. N is the number of people of the travelling group, which is a crucial variable when choosing private car transport. a represents vehicle operating costs (€/km) and depends on the type of vehicle and consequently on the different user profiles (business, economy and family). b is the value of time (€/h) and c is the tariff perception (%).

The seven variables can be grouped into those that remain constant on a single route (D, T, P) and those that vary according to the demand segment considered (a, b, c, N).

2.3.2. *Public transport tariffs and generalised cost formula*

In the case of collective transport, the generalised cost formula becomes

$$G_C = (bT + cP)N.$$

The main differences between this formula and the one associated with private transport is related to the role of variable N (number of persons in the group) and to the definition of fares P.

Table 2. Main characteristics assumed in the simulation for each demand segment.

	Operating cost	Value of time	Tariff perception	N. person per vehicle
Business	High	High	Medium	1
Economy	n.a.	Low	High	1
Family	Medium	Low	Medium	3

In the private vehicle, the variable N influences only the value of trip time component; to the contrary, in public transport, it multiplies the entire generalised cost.

The general formula used to describe the fare for each O/D relation is

$$P = p_0 + pd,$$

where P is the univocal price/tariff of a specific route, made of two components: d is the distance and p is a component proportional to distance, plus a fixed component independent from distance p_0. This formula allows one to reproduce either tariffs perfectly proportional to the distance travelled or tariff perfectly flat (for instance, tariff of low-cost flight or of local public transport services). The parameters p_0 and p were calculated on the basis of real tariffs extrapolated from transport operator websites. Both depend on the transport mode, on the type of service (train class, low-cost or full fare airline and coaches applying yield management or not), moment of purchase and also the presence of competition on the route. For example, we used different fares for high-speed trains running on the Milan–Rome route, where some airlines and two rail operators compete, and on the Milan–Venice route, where Trenitalia is the only operator. The same happens for air prices.

2.4. Construction of the hypergraph and mapping

The main difference between a transport graph and a *hypergraph* is that the second one also considers the temporal dimension in addition to the spatial variable. Each node is therefore representative of both a particular place (e.g. a station) and a specific time of the day (for instance, the arrival time of a single train) and takes the name of *hypernode* (Camus & Rupi, 2001; Cascetta, 1990; Gallo, Longo, Pallottino, & Nguyen, 1993). Therefore, a link (or, better, a *hyperlink*) connects two different times of day, in different places or even in the same place (i.e. dwell time in a station) while it has a cost equal to the difference between the times of two nodes plus the tariff, if any. Moreover, unlike the infrastructure graph, the hyperlinks are necessarily unidirectional since the possible movements are only the ones with increasing time.

We created a dense network of connectors that allows us to model the links between the logical structure described above and both the infrastructural graph and the sub-provincial zoning. These connectors can be grouped in three categories. The *main connector* provides the link between the centroid of a catchment area and a node of the graph corresponding to a stop/station. The *departure connector* provides the link between the stop/station and a time of the day inside that station. The *arrival connector* represents the time of the day in which users arrive in that specific stop (Figure 5).

With this architecture, the *hypergraph* is able to reproduce even complex travel choices based on multiple carriers (for instance, coach–train interchange or train–flight interchange). Therefore, the entire timetable database is structured to interface with the *hypergraph* of transport services and, using a minimum paths algorithm (Dijkstra, 1959), is able to return the

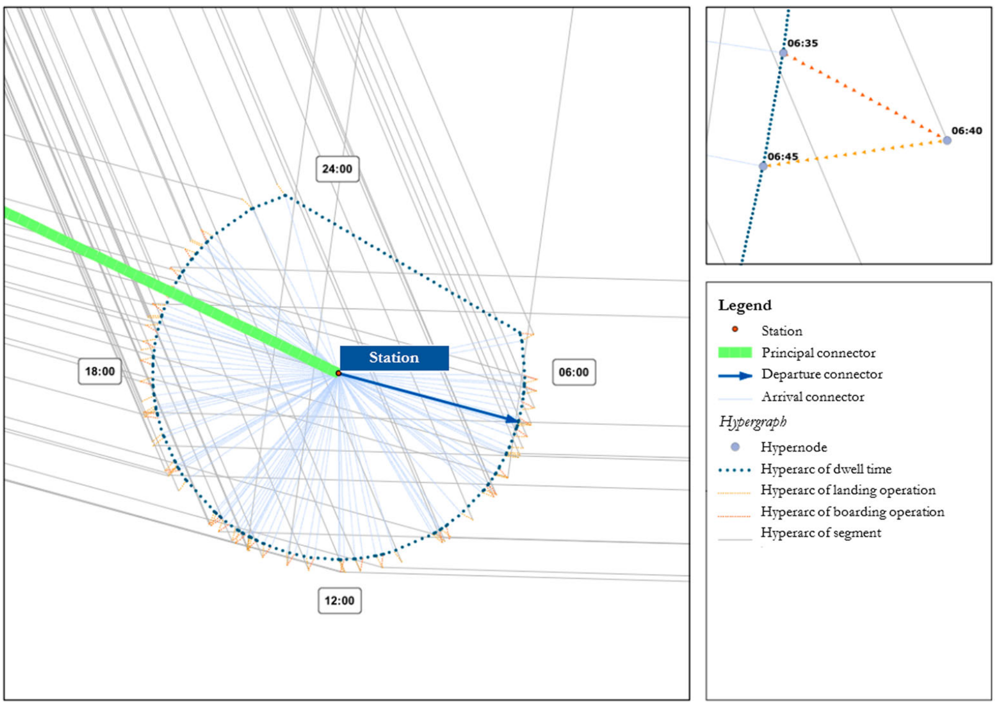

Figure 5. Graphic showing the model structure used to calculate real interchanges on the base of public service timetables.

different relationships among the 371 areas with the different modes of transport.

Thanks to this tool, we are able to calculate and map, for each demand segment and mean of transport, a generalised cost that assumes, as main variable, the time of the day in which a single trip takes place. Considering long-distance O/D relations, time variables, especially for public services, play a fundamental role because in most cases, to complete a journey, an interchange is needed.

In this regard, the use of a hypergraph (real service timetables) allows to exclude, from the calculation of the minimum path algorithm, those services which, despite standing on the same interchange node, are not aligned with the time of departure/arrival of the public service in that specific stop/station. Moreover, destinations located at the same distance from the starting point could present heterogeneous generalised costs due to the fact that waiting time at the interchange node may be very different.

The Main Map produced includes two different representations, both derived from the timetables database: the charts of the supply for the three transport modes (rail, coach and air) and the charts of the generalised costs calculated from two selected origins for one demand segment. Generalised cost charts include both mono-modal and multimodal simulations, assumed to depart at 7 am.

3. Conclusions

The Main Map presented here shows the 2013–2014 distribution of national medium and long-distance public transport services and the model behind it allows calculating generalised costs for each O/D relation and mode. The Main Map is based on the described model and thus includes and integrates both the space and time component of the trips.

This comprehensive and complete representation of all long-distance domestic transport in Italy is new and it never existed before. Apart from the opportunity to visualise the entire network of services, it allows to derive some relevant facts related to the Italian transport system.

The coach services, as the chart clearly shows, are structured on the paths of post-Second World War migrations from the South of Italy to the industrial and service cities located in the Centre-North/West (e.g. Rome, Milan and Turin). At the same time, this historical structure overlooks areas with more recent high market potential such as the foothill areas of North-eastern Italy.

The air service is strongly oriented along the North–South axis and around the Alitalia hub of Rome. The route Fiumicino–Linate remains, despite the recent shift to high-speed rail, the one with the largest number of flights per day. Its highest number of domestic flights confirms the role of Fiumicino airport as the main domestic hub.

The rail service is increasingly concentrated along the main routes (Milan/Venice–Rome–Naples axis, but also Milan–Genova, Turin–Venice and Bologna–Bari), with the node of Bologna having the highest number of trains per day. Other secondary lines tend to have much less trains per day, except those around main metropolitan areas where a relevant regional traffic exists for a few dozen kilometre.

The generalised costs charts visualise in a clear way the different markets of the transport modes. For example, air services are the cheapest way to reach the areas around southern Italian cities such as Catania or Cagliari from Milan or Rome, while the rest of southern provinces remain easier to be reached by train. Coaches are seldom the cheapest way to move (because slower), but their generalised cost differential is sometimes very small and, in fact, they keep a steady or increasing market share in the economy segment, also thanks to the fact that they often do not need interchanges. The rail charts present some time–space 'irregularities': for rail services, the distance is less and less the main variable defining travel time and the speed of the line warps the generalised cost areas. For example, the new high-speed line makes the costs to reach Naples from Milan (approx. 800 km) cheaper than to reach Trieste (400 km) or the Adriatic coast. Similarly, marginal mountain areas, where rail is absent or weak, show very high travel cost also for limited distances.

Interestingly, both for Milan and Rome, the rail and multimodal generalised cost charts are matching in many parts. This means that the rail mode is often the most affordable, also in comparison with multimodal options, which become 'visible' only when reaching the southern remote areas. This is due to the role of Milan and Rome as major nodes of the Italian railway system, but also gives evidence to the fact that multimodal long-distance transport (excluding local transport) is still undeveloped in Italy or limited to few air–rail connections.

Acknowledgements

The authors wish to give thanks to some of the people who helped us to build parts of the database: Ila Maltese, Martina Ferro, Soorush Baghban, Giulia Musso, Maryam Chegeni, Jonata Di Benedetto and Simone Borghi.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Paolo Beria http://orcid.org/0000-0001-5171-6576

Raffaele Grimaldi http://orcid.org/0000-0002-7832-2398
Antonio Laurino http://orcid.org/0000-0003-3803-8066
Alberto Bertolin http://orcid.org/0000-0003-3804-3551

Software

Microsoft Excel 2013 and Microsoft Access 2013 were used to standardise and elaborate input data. ESRI ArcGIS 10.2 was used both as the analysis platform for this project and to create the output maps. Interface processes are mainly developed in Python 2.6.

References

Assaeroporti. (2014). *Associazione Italiana Gestori Aeroporti*. Retrieved September 22, 2014, from http://www.assaeroporti.com

Boitani, A., & Ponti, M. (2006). Infrastrutture e politica dei trasporti. *il Mulino*, *55*(1), 102–112. doi:10.1402/21261

Camus, R., & Rupi, F. (2001). Modelli di fermata e calcolo degli ipercammini minimi nelle reti di trasporto pubblico urbano. In C. Podesta & R. Maja (Eds.), *Metodi e modelli per il trasporto pubblico locale* (pp. 147–160). Milano: F. Angeli.

Cascetta, E. (1990). *Metodi quantitativi per la pianificazione dei sistemi di trasporto*. Padova: CEDAM.

Cascetta, E. (2006). *Modelli per i sistemi di trasporto: teoria e applicazioni*. Torino: UTET università.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, *1*, 269–271.

ENAC. (2014). *Italian Civil Aviation Authority*. Retrieved September 22, 2014, from http://www.enac.gov.it

European Commission. (2007). *Regions in the European Union – Nomenclature of territorial units for statistics – NUTS 2006 /EU27*. Luxembourg: Office for Official Publications of the European Communities.

Gallo, G., Longo, G., Pallottino, S., & Nguyen, S. (1993). Directed hypergraphst and applications. *Discrete Applied Mathematics*, *42*(2–3), 177–201.

ISTAT. (2014). *Istituto Nazionale di Statistica*. Retrieved September 22, 2014, from http://www.istat.it

Ministero Delle Infrastrutture e dei Trasporti. (2014). *Conto Nazionale delle Infrastrutture e dei Trasporti – Anni 2012–2013*. Roma: Istituto Poligrafico e Zecca dello Stato.

Nichols, A. J. (1975). *Standard generalized cost parameters for modelling interurban traffic and evaluating interurban road schemes*, Note 255, London: Mathematical Advisory Unit, Department of the Environment.

Ortúzar, J., & Willumsen, L. G. (1990). *Modelling transport*. Chichester: John Wiley & Sons.

Zofío, J. L., Condeço-Melhorado, A. M., Maroto-Sánchez, A., & Gutiérrez, J. (2014). Generalized transport costs and index numbers: A geographical analysis of economic and infrastructure fundamentals. *Transportation Research Part A: Policy and Practice*, *67*, 141–157.