http://www.istat.it
Contact Centre
Press Office
tel. +39 06 4673.2243/4
ufficiostampa@istat.it

THE ACCESSIBILITY OF MUNICIPALITIES TO MAJOR TRANSPORT INFRASTRUCTURES – Methodological notes

Year 2022

Criteria for selecting point infrastructures

Railway stations with active passenger service

The total number of active railway stations where trains provide passenger service is 2,075. The source of the data is RFI (Rete Ferroviaria Italiana - Italian Railway Network). Considering that air and sea travel are only partially related to work reasons and mostly pertain to non-short routes (see further sections), for consistency, information on the number of distinct trains per day based on distance has been used, i.e., regional or long-distance trains. The stations where long-distance trains make stops are 283, as represented in the figure below:

FIGURE 1. RAILWAY STATIONS WITH ACTIVE PASSENGER SERVICE AND STOPS FOR LONG-DISTANCE TRAINS. YEAR 2022

In nine cases, including Napoli Afragola, Bologna/AV and Reggio Emilia AV, which are noteworthy, the stations have only long-distance train traffic and zero regional trains, while in the remaining 274 stations, regional train traffic is always present. Out of the 283 stations, 100 have long-distance train traffic of at most 5 trains per day; 146 have at most 10; 186 have at most 15. Given the heterogeneity, it is appropriate to establish a threshold of trains per day based on which to choose the reference universe for calculating accessibility indicators.

To make this selection as less arbitrary as possible, reference was made to the Zipf distribution, the counterpart of the Gaussian distribution in various domains. In short, in distributions that satisfy Zipf's law, the rank occupied by individual stations - in our case, the subset with long-distance trains - has a well-defined relationship with the number of trains per day: the number of trains decreases approximately with the rank of the variable itself. Therefore, the station that ranks first in terms of the number of trains per day has twice as many trains as the second station, which in turn has twice as many trains as the third station, and so on. The law can be characterized by a linear relationship between the logarithms of rank and probability or frequency. By applying this statistical rule, it was verified for which minimum number of trains per day the above-mentioned linear relationship is most respected (where the parameter β is closest to 1 in absolute value). This led to the exclusion of stations that have a number of trains below this threshold.

Based on the joint reading of graphs and estimation parameters, the optimal threshold seems to be trains per day > 3, for which β has a value of -1.04 and a higher R^2 value than the previous threshold (trains per day > 2), as evidenced by the figures below::

FIGURE 2. RELATIONSHIP BETWEEN THE NUMBER OF TRAINS (IN LOG) AND THE RANK OF THE STATIONS (IN LOG)).

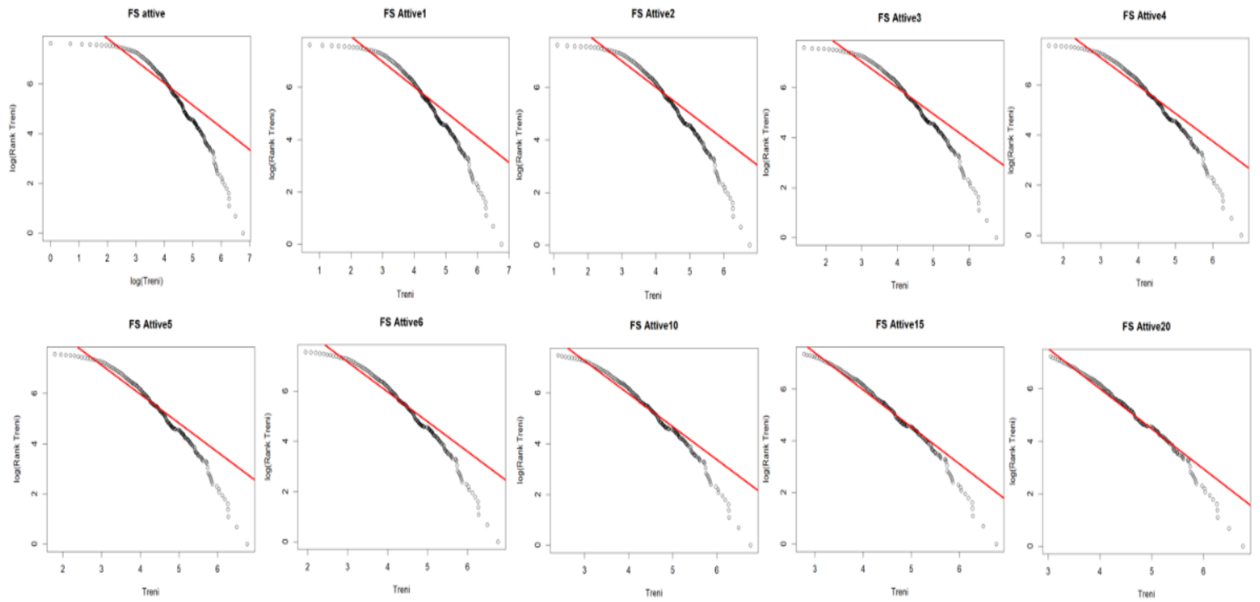
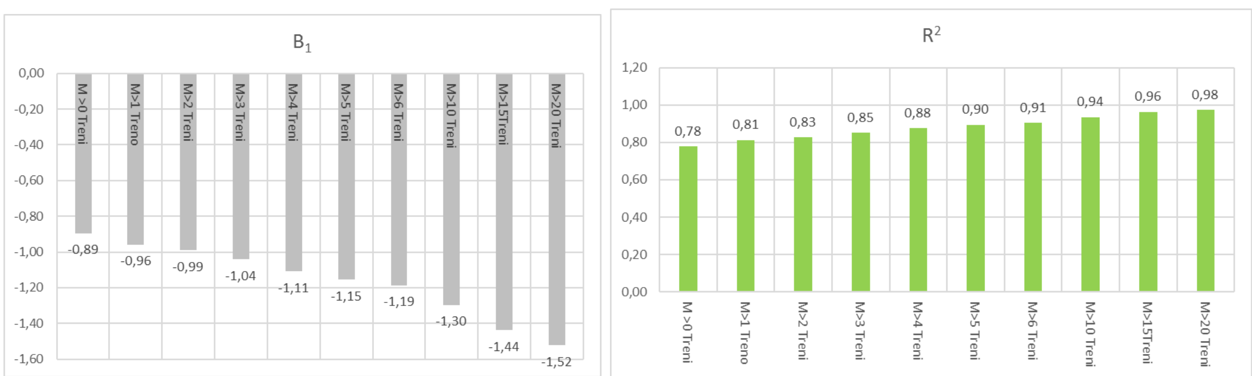


FIGURE 3. ESTIMATED COEFFICIENTS B AND R^2 OF THE MODELS



A particular case is represented by Sardinia, which does not have long-distance trains for obvious reasons and has transit limited to regional trains only. The same statistical rule was applied to regional trains, resulting in an optimal threshold of >5 trains per day. By applying these thresholds, ultimately, a total of 258 railway stations with active passenger service were considered for Italy as a whole.

Access to the highway network (expressways for Sardinia)

Access to the highway network was calculated by considering all access points to the highway artery, including access ramps (2,660+182 for Sardinia). The calculations were carried out from all Italian municipalities to all access points. Access to the highway network is provided in the TomTom graph as road POIs (Points of Interest) and are in UTM-WGS84 geographic coordinates.

Expressways for the region of Sardinia are those similar to those identified in the TomTom graph with a road travel speed of at least 90 km/h and classified as main rural roads. The reference universe was limited to:

- State road 131 Carlo Felice (SS 131), which is the main road artery of Sardinia, connecting the island from north to south from Cagliari to Porto Torres;

- State road 729 Sassari-Olbia (SS 729), an Italian state road, still partially under construction, connecting the northwestern area of Sardinia with the eastern coast near Olbia;

- State road 130 Iglesiente (SS 130), connecting Cagliari with Iglesias;

- SS291 connecting Sassari to Alghero;

- SS554 Cagliaritana, an Italian state road of local relevance, connecting the northern part of the hinterland of Cagliari;

- SS195 var Sulcitana, a section connecting Cagliari to Giorgino.

Access points to these main Sardinian rural roads were identified through careful examination of the road route using Google Maps and GIS tools, also utilizing road network data provided by TomTom. From the Google Maps graph containing all access ramps to the territory served by these roads, the coordinates of the ramps were extracted to make the data homogeneous with that provided by TomTom, where access points to the highway network are provided with UTM-WGS84 geographic coordinates.

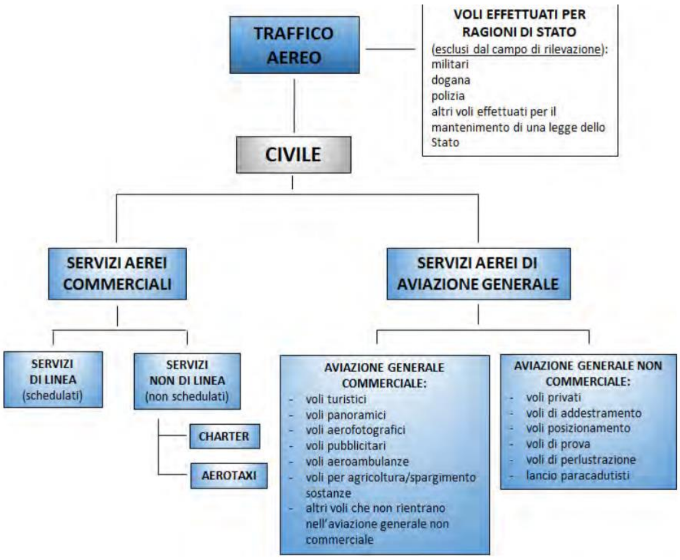
The coordinates of the access points obtained through Google Maps were converted into UTM-WGS84 geographic coordinates to enable distance calculation in ArcGIS.

Distances to access points to the main Sardinian arteries were determined using the same methods as for access to toll booths, i.e., from all municipalities in the region of Sardinia to all access points in the region.

Airports

For airports, the initial reference was the ENAC (Italian Civil Aviation Authority) observation field regarding civilian air traffic, as reported below:

FIGURE 4. OBSERVATION FIELD FOR AIRPORT SELECTION.



Source: Enac

Some considerations that led to the final selection:

- The focus is on commercial air services.
- The ENAC Traffic Data Report considers 45 airports. However, the official list on the website includes only 39 of national interest (https://www.enac.gov.it/aeroporti/infrastruttureaeroportuali/aeroporti-in-italia).
- When is an airport of national interest? "In accordance with the criteria set out in Article 698 of the Navigation Code, airports and airport systems of national interest are identified as essential nodes for the exercise of the State's exclusive competences, for each of the ten traffic basins identified in the national territorial network, as specified below and in compliance with the conditions set out in paragraphs 4, 5, and 6." Source: PRESIDENTIAL DECREE OF THE REPUBLIC of September 17, 2015, No. 201, Regulation identifying airports of national interest, pursuant to Article 698 of the Navigation Code. GU General Series no. 294 of 12-18-2015. The selection takes into account the size and type of traffic, the territorial location, and the strategic role of the airports, as well as what is provided for in the TEN European projects. The measure is aimed at developing the sector within a governance framework that balances the needs of national and international traffic demand with the development needs of the territories, the enhancement of necessary infrastructure, the effective use of public resources, and the efficiency of air navigation services and other services provided within the airport context.

Of the 39 airports of national interest, 4 were eliminated because the passenger traffic in them is either null or very low. Specifically, these are: Foggia/Gino Lisa, Taranto/Grottaglie, Salerno/Pontecagnano, and Brescia/Montichiari.

The final universe consists of 35 airports. All the airports considered have national passenger traffic, and almost all of them also have international passenger traffic, as evidenced by ENAC publications on air traffic data (https://www.enac.gov.it/aeroporti/infrastrutture-aeroportuali/dati-di-traffico).

Ports with passenger service

The starting universe for the selection of ports considered in the accessibility analysis is that of statistical ports, used by Istat (source: Istat, Maritime Transport Survey) for the dissemination of data on passenger traffic, goods (with details on the type of cargo for major ports, which transport more than 1 million tons of goods per year), and ships, defined in accordance with Eurostat.

A "port," for statistical purposes, is an area of land and water equipped with infrastructure and equipment suitable primarily for the mooring of vessels and the carrying out of related cargo handling, loading, and unloading operations, receiving and delivering goods, embarking and disembarking passengers, crew members, and other individuals, and any other infrastructure necessary for transport operators within the port area.

Specifically, the "statistical port" includes one or more ports, normally controlled by a single port authority or port captaincy, capable of recording ship movements, passenger transfers, and cargo transfers. It includes "commercial" seaports, i.e., ports with access to the sea that have commercial, logistic, industrial, and oil traffic of goods and/or passengers, including cruise passengers. It excludes river ports and those primarily or exclusively for tourism and/or fishing purposes.

In total, in 2022, there were 143 statistical ports, including major and minor ones. In many cases, individual port data is aggregated with that of the main port. Therefore, we chose to refer to the narrower universe consisting of only the main ("reference") port compared to those grouped together.

This reduced the number of ports to 79 (excluding the Tremiti Islands port). In addition to this selection, it was deemed appropriate to further refine the choice of ports by considering the passenger traffic recorded in the year. The selected threshold is 10,000 passengers per year. Ultimately, our universe consists of 54 ports with passenger traffic.

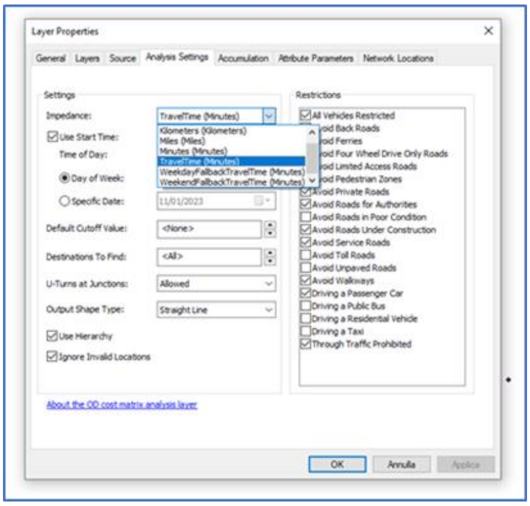
Distance matrices and travel time calculations

The calculation takes the municipality's centroid as a reference point, specifically the centroid of the census track containing the Municipality's Town Hall, and refers to Italian municipality data as of January 1, 2021 (7,903 municipalities). For the Municipality of Rome, only the centroids of the 15 boroughs were considered. For reference infrastructures, centroids generated from the coordinates of the infrastructure's representative point provided by the relevant authorities were used (see previous sections for details).

Among the solvers available in the Network Analyst extension, the OD Cost Matrix solver was chosen, which finds and measures routes along a network (graph) from multiple origins to multiple destinations based on the well-known Dijkstra's algorithm.

Through the computations, an OD cost matrix analysis layer is created, appearing in the Network Analyst window along with its six network analysis classes: Origins, Destinations, Lines, Point Barriers, Line Barriers, and Polygon Barriers. Analysis parameters are set in the layer properties dialog box for all analysis layers, as illustrated in the image related to the layer properties:

FIGURE 5. LAYER PROPERTIES OF NETWORK ANALYST



Impedance represents a cost attribute. For example, Travel Time evaluates traffic taking into account traffic speed for a specific time of day, specifying a start time and date (or day). When providing a start time, the resulting cost matrix considers the varying traffic speeds at that time and on that date.

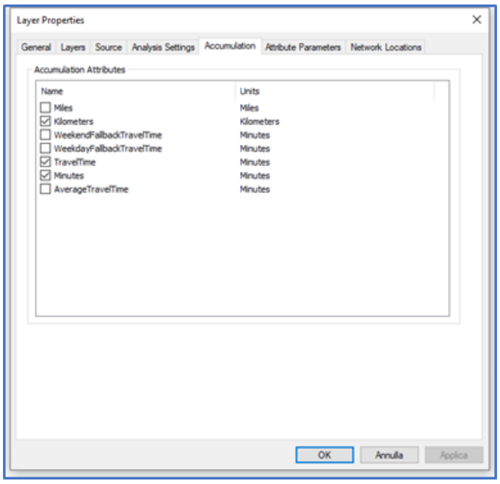
Travel times are calculated based on the speeds of individual road segments called functional road classes, totaling eight in all. Restrictions are attributes that limit the calculation of destinations.

It is possible to choose which restriction attributes must be respected during the analysis resolution. In most cases, restrictions may prohibit, avoid, or prefer certain roads. A restriction attribute, such as Oneway, is used when finding solutions for vehicles that must adhere to one-way streets (e.g., non-emergency vehicles). Among the restriction attributes are those that consider one-way streets (oneway) and those that consider height or weight limits or the passage of hazardous materials. Additionally, it can be defined whether to allow U-turns anywhere, nowhere, only at intersections and dead ends.

OD-cost-matrix cost matrices can be represented with linear geometry or even without geometry. In both cases, the path is still calculated along the network and returns the same total cost in the attribute table as a result of path analysis.

Hierarchy is the order or rank assigned to elements of the network. A road network has an attribute on origin characteristics that divides roads into three (or more) classes, such as local, secondary, and primary.

FIGURE 6. COST ATTRIBUTES SELECTABLE BY NETWORK ANALYST

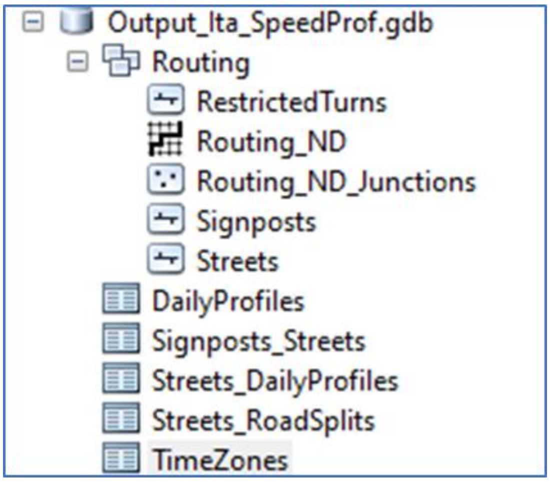


Other elements that can be chosen are related to those cost attributes that can be selected from the network dataset to accumulate on line objects, which represent the lowest-cost paths along the network.

For each accumulated cost attribute, a Total_[Impedance] property is added to the rows output by the solver, where [Impedance] is replaced with the name of the accumulated cost attribute.

The road graph used is accompanied by speed profiles provided by TomTom, which contain historical traffic data. This information can be transferred to the dataset to generate the best routes.

FIGURE 7. SELECTABLE SPEED PROFILES FOR GENERATING THE BEST ROUTES.



The data is processed at the municipal level and contains information about routes from all municipalities to all considered infrastructures. The route of movements is calculated in only one direction, from municipality A to infrastructure A.

Logic of the accessibility-proximity model

Structure of a generic accessibility indicator

Accessibility is a fundamental concept, the meaning of which is known to all but often difficult to define until the indicator used to measure it is specified. In this study, accessibility to specific points is analyzed, rather than evaluating the level of accessibility of an area. For example, if one wanted to assess the level of accessibility of a municipality, the quantity of transportation services present within it could be measured, such as kilometers of highways or the number of railway stations. This concept of accessibility is linked to infrastructure provision.

In general, to consider an accessibility indicator, it is necessary to take into account several dimensions, which are specified in relation to the analysis under consideration:

TABLE 1. DIMENSIONS TO BE CONSIDERED

DIMENSIONS	DESCRIPTIONS
Origins	We use Municipalities centroids (year 2021) based on the Town Hall address
Destinations	Location of infrastructure, opportunities sought in the area
Transport modality	Only car transport is considered
Impedence	Time travelled by car calculated using Tom-Tom's speed profiles

To understand the level of accessibility of a municipality in relation to various infrastructures, we proceed by calculating the time required to reach the nearest infrastructure, which constitutes an indicator of cost in terms of the time needed to travel to the infrastructure.

In general, we define the level of accessibility of a municipality as a generic function of the form:

$$A_i = \sum_j g(W_j) f(c_{ij})$$

Where

- A_i represents the accessibility level of the origin point *i*;
- *j* denotes the different infrastructures;
- $g(W_j)$ is a function that evaluates the importance of infrastructure *j* based on certain criteria;
- $f(c_{ij})$ is a function that calculates the cost, in this case the time, to reach infrastructure *j* from point *i*.

The accessibility level A_i for the origin point *i* will be calculated with respect to an area of interest defined by the opportunities that are considered relevant for this municipality. Thus, the indicator will be relative to the j opportunities that can be considered. Examples of such situations include considering the first *n* reachable opportunities or the opportunities within the region of belonging.

Let's consider a generic accessibility indicator for municipalities towards infrastructures, in a cumulative opportunity form, which will be of the type:

$$\begin{cases} A_i = 1 \; if \; c_{ij} \leq \bar{c} \\ A_i = 0 \; if \; c_{ij} > \bar{c} \end{cases}$$

Given the maximum acceptable cost level, denoted as $\bar{c}$, our municipality will be deemed accessible $A_i = 1)$ if it is able to reach at least one opportunity within the maximum value; otherwise, it will not have a sufficient level of accessibility.

This type of analysis conforms to what is used for inland areas and naturally depends on the threshold level, which is arbitrary. However, with the same chosen value, a municipality will be more or less accessible depending on:

$$g(W_i) = f(transport\ modality,\ network,\ spatial\ distribution\ of\ W_i)$$

Indeed, the number of reachable opportunities depends on the mode of transportation used. For instance, if we consider accessibility to an infrastructure within 30 minutes, the opportunities reachable by foot will inevitably be fewer than those reachable by car.

With the same mode of transportation, the road network affects travel times. Consider areas without highways, which will inevitably have lower travel speeds, impacting the accessibility levels for the accessibility indicator.

Nevertheless, under the same conditions of the c_{ij} parameter, the spatial distribution of opportunities significantly influences defining the shortest route. Consider the accessibility of a municipality located in the province of Bolzano concerning ports and railway stations. The fact that the spatial distribution of ports is tied to the sea, while railway stations are more dispersed across national territory, implies that the municipality is less accessible to port infrastructure, despite the presence of high-traffic roads.

The impact of the spatial distribution of opportunities

Is it possible to measure the impact of the spatial distribution of opportunities concerning the accessibility indicator?

A recent OECD study proposes a suitable methodology to understand the phenomenon. It starts with the juxtaposition of two aspects:

• Absolute accessibility: the ability to reach a location within a certain maximum time in a specific mode

• Proximity: the presence of opportunities within a predetermined linear distance.

Internationally, 8 km of linear distance is considered comparable to a 30-minute car journey. To better understand this situation, consider the comparison in Figure 8 for an ideal municipality i represented at its centroid (blue point).

Starting from the municipality's centroid, the 30-minute isochrone is calculated, representing the maximum area reachable within this time limit, and compared with the circle of radius 8 km. Regarding the examined origin point (blue point), the four represented opportunities have different characteristics in terms of accessibility and proximity (Figure 9).

FIGURE 8. COMPARISON OF ABSOLUTE ACCESSIBILITY AND PROXIMITY.

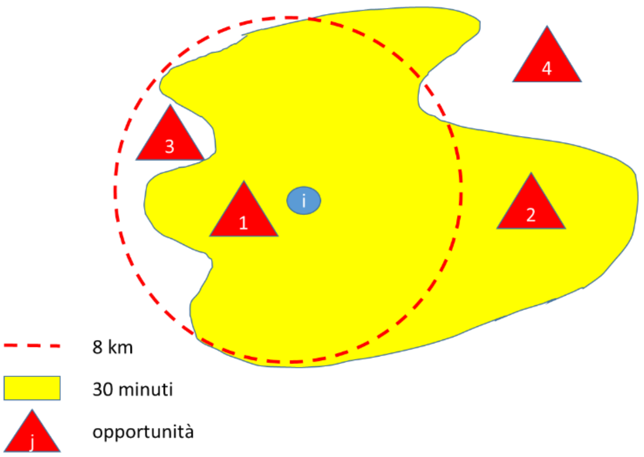


FIGURE 9. LOGICAL CLASSIFICATION OF ACCESSIBILITY-PROXIMITY COMBINATIONS.

		Proximity	
		YES	NO
Absolute accessibility	YES	1-1 Accessible and Nearby	1-2 Accessible but Distant
	NO	2-1 Inaccessible but Nearby	2-2 Inaccessible and Distant

The identifiable cases can be summarized as follows:

1. Accessible and nearby opportunities, reachable within 30 minutes and within a straight-line distance of less than 8 km (cluster 1-1);
2. Accessible opportunities despite not being nearby, destinations benefiting from a particularly fast road network allowing timely access even though they are linearly farther than the considered limit (cluster 1-2)
3. Inaccessible opportunities even if nearby, falling within the 8 km range but not reachable within 30 minutes (cluster 2-1);
4. Inaccessible opportunities because not nearby, a typical situation for all destinations beyond the 8 km range and therefore rightly inaccessible (cluster 2-2).

The objective of this analysis is to compare the level of accessibility for Italian municipalities concerning major transport infrastructures. The significant diversity in the distribution of these infrastructures necessitates a revision of reference limits. Specifically, situations such as ports, within a national context, would render the majority of Italian municipalities inaccessible because the 30-minute and 8 km linear limits would be too stringent. Therefore, it was decided to find an endogenous criterion for distributions so that reality would indicate the most appropriate limit. On the other hand, there is no regulatory reference ensuring the reaching of infrastructures within a limit.

The natural reference has become the median of the distributions of transport times to the nearest infrastructure and the linear distance to the nearest infrastructure. Obviously, this choice means that half of the municipalities are in a good situation of accessibility and proximity, but it is a way to begin addressing the reflection on the topic.

Selection of the cost function and its parameters in the gravitational model

The accessibility index, introduced by Hansen (1959), and borrowed on the basis of the gravitational law, can be described by the following formula:

$$A_i = \sum_j W_j * f(c_{ij})$$

where:

- A represents the measure of accessibility;
- i indicates the municipality;
- j indicates the infrastructure present in an area;
- W indicates the supply coming from infrastructure j;
- cij is the cost incurred to reach infrastructure j from municipality I;

• $f(cij)$ finally represents the decay function, with a cost in terms of time required to reach an infrastructure that increases with the increasing distance between municipality and infrastructure.

Such decay function can be rewritten as:

$$f\left(c_{ij}\right) = f(d_{ij}, \lambda),$$

In which, dij represents the distance in minutes between municipality i and infrastructure j, λ is a parameter, a sort of friction cost, which assigns a greater penalty to locations farther away compared to those closer.

In summary, the indicator Ai represents the quantity of accessibility services potentially reachable from that point and conceptually equals a weighted average of the number of n considered infrastructures for the cost required to reach them.

Regarding the services offered by the infrastructure (W), the following information has been used:

- For train stations: the number of trains per day serving passengers at the selected stations.
- For access to the highway network, it is assumed that there is no difference between them in terms of supply.
- For airports: the number of aircraft movements generated.
- For passenger ports: the number of ships docking annually at a given port.

Regarding the cost function (also known as decay or impedance function), to account for the fact that the willingness to reach the various infrastructures under analysis changes based on their relative distance (but also based on the reason for the travel), it was chosen - following recent literature (Palacios and Le Geneidy, 2022; Bauer and Groneberg, 2016; Luo et al., 2014; Wan et al., 2012; Kwan, 1998) - to use a non-linear function, $f(cij)$, of negative exponential type. "Several studies use different impedance functions, such as power, Gaussian or logistic functions; however, the negative exponential function is the most often used and also the most closely tied to travel behavior theory (Karst T. Geursa and Bert van Wee, 2004)".

The use of such a function type penalizes municipalities farther away from the infrastructures, creating a sort of discouragement effect, where the attractiveness capacity of services WJ decreases more than proportionally with the increase in distance. Practically, it means that either one has an infrastructure nearby, or it is not used. For this reason, the function "decays" beyond a certain distance.

This type of function, as mentioned, is widely used in literature (Palacios and Le Geneidy, 2022; Bauer and Groneberg, 2016; Luo et al., 2014; Wan et al., 2012; Kwan, 1998) and can be analytically described by the following formula:

$$f(d_{ij},\lambda) = e^{-\lambda d}$$

Where lambda (λ) represents a relevant parameter for understanding the speed at which travel availability decreases as the time taken to reach the infrastructure (d) increases. An individual's willingness to reach a certain infrastructure decreases as the time required to reach it increases.

Therefore, the choice of the parameter $\boldsymbol{\lambda}$ becomes relevant. Usually, this parameter is estimated based on the availability of data from surveys (or other sources).[1]

[1] Stat Canada (Alasia, Alessandro, Frédéric Bédard, Julie Bélanger, Eric Guimond, and Christopher Penney. 2017. "Measuring remoteness and accessibility: A set of indices for Canadian communities." Reports on Special Business Projects, Statistics Canada), uses the following distribution function on the basis of the survey data at its disposal: $f(d_{ij},\lambda) = 1 - (1 - e^{-\frac{d_{i,j}}{\lambda}})$ such that the parameter lambda, based on survey data, is chosen so that the median duration of the willingness to travel to a certain infrastructure corresponds to the midpoint of the interval [0, 1].

In the figure below, you can observe the different trends taken by some examples of impedance functions[2] with some of the different parameters used in the literature[3]. The functions closer to the origin tend to penalize municipalities farther away from the infrastructures more and thus reward proximity to an infrastructure in terms of accessibility.

In our specific case, the absence of specific empirical data through which to measure people's willingness to travel (to cover a certain distance) to reach a particular infrastructure complicates the analysis and requires us to make assumptions about the impedance function. In fact, the absence of data related to the behavioral choices of the populations of individual Italian Municipalities forces us to conform to what has emerged in the literature regarding studies conducted in other countries. This implies assuming that Italians' willingness to travel to reach an infrastructure is identical to what has been observed in other countries and/or that the "conditions" of the infrastructures in different countries are similar and/or that the configuration of places and their related travel costs are similar.

Additionally, in this analysis, we are proceeding with a territorial comparison, while the choices are individual, and therefore, we should consider the motivations for travel (commuters, tourism, for example) that significantly influence the willingness to incur initial travel costs.

The only empirical support available to try to estimate the impedance function and the related parameter "λ" comes from data related to the commuting matrix of 2011. Through this matrix, it is possible to make assumptions regarding the railway transport infrastructure. With this data, it is indeed possible to analyze the relationship between the flow of commuter individuals who choose to use the train to reach their workplace or study place in each municipality and the distance that municipality has from the nearest railway station. Or, as in the case of highway toll booths, the relationship between the flow of commuters using the car to travel from one municipality to another, taking more than 30 minutes.

---

2 Inverse Function: $f(d_{ij},\lambda) = d^{-1}$; Inverse Square Function: $f(d_{ij},\lambda) = d^{-2}$; Gaussian Function: $f(d_{ij},\lambda) = e^{\frac{-(d^2)}{1000}}$; Negative Exponential Function: $f(d_{ij},\lambda) = e^{-0,08*d}$; Stat-Canada fuction (median 30'): $f(d_{ij},\lambda) = 1-(1-e^{-\frac{d_{i,j}}{43}})$; Stat-Canada Function (median 10'): $f(d_{ij},\lambda) = 1-(1-e^{-\frac{d_{i,j}}{14}})$

3 For more details, see the following works:

Alasia et al., Alessandro, Frédéric Bédard, Julie Bélanger, Eric Guimond, and Christopher Penney. 2017. "Measuring remoteness and accessibility: A set of indices for Canadian communities." Reports on Special Business Projects, Statistics Canada

Apparicio, P., & Seguin, A.-M. (2006). "Measuring the Accessibility of Services and Facilities for Residents of Public Housing in Montreal". Urban Studies, 43(1), 187-211. https://doi.org/10.1080/00420980500409334;

Bauer J, Groneberg DA (2016) Measuring Spatial Accessibility of Health Care Providers – Introduction of a Variable Distance Decay Function within the Floating Catchment Area (FCA) Method. PLoS ONE 11(7): e0159148. https://doi.org/10.1371/journal.pone.0159148.

Geurs KT, B Van Wee (2004). "Accessibility evaluation of land-use and transport strategies: review and research directions", Journal of Transport geography 12 (2), 127-140;

Hansen, W. G. (1959). How accessibility shapes land use. Journal of the American Institute of Planners, 25(2), 73–76;

Kapatsila, Bogdan & Palacios, Manuel Santana & Grisé, Emily & El-Geneidy, Ahmed, 2023. "Resolving the accessibility dilemma: Comparing cumulative and gravity-based measures of accessibility in eight Canadian cities," Journal of Transport Geography, Elsevier, vol. 107(C).;

Kwan, M. P. (1998). Space-time and integral measures of individual accessibility: a comparative analysis using a point-based framework. Geographical Analysis, 30(3), 191-216. https://doi.org/10.1111/j.1538-4632.1998.tb00396.x

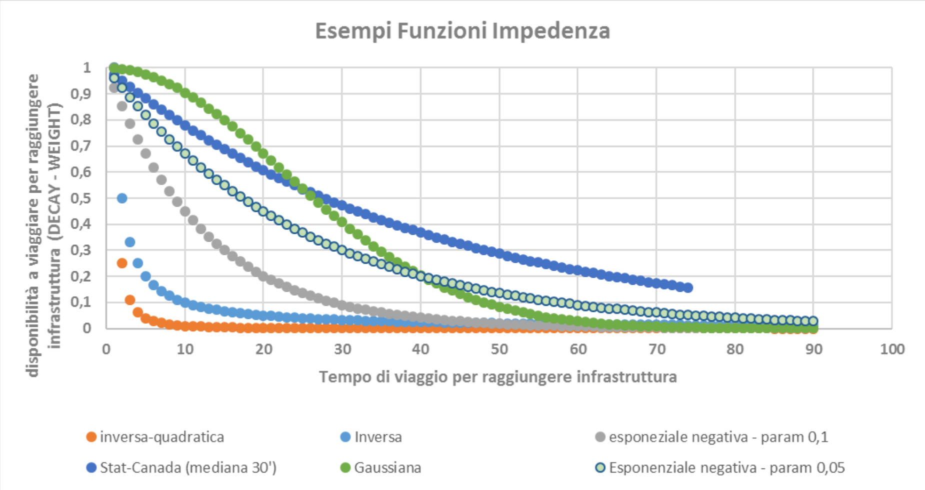
Liu, S., & Zhu, X. (2004). Accessibility analyst: An integrated GIS tool for accessibility analysis in urban transportation planning. Environment and Planning B, 31(1), 105–124.

Luo W, Wang F. (2003) Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. Environ Plann B.;30:865–84

Luo, W. (2004). Using a GIS-based floating catchment method to assess areas with shortage of physicians. Health & Place, 10(1), 1–11.

Luo, J. (2014). Integrating the Huff model and floating catchment area methods to analyze spatial access to healthcare services. Transactions in GIS, 18(3), 436–448.

FIGURE 10. EXAMPLES OF IMPEDANCE FUNCTIONS



Observing this data, it is evident that with the increase in the distance of the municipality from the railway station or highway toll booth, the number of individuals who choose to travel by train or car as a means of transportation decreases. This reduction appears to occur nonlinearly, more akin to a negative exponential function (see graph below). In Figure 4.2, individual points represent the population of municipalities (grouped in ranges of five minutes distance from the nearest railway infrastructure) willing to choose the train to reach their workplace or study place, and the distance in ranges of 5 minutes.

To model the relationship between distance and commuter flow, we adopted a Poisson regression. Using the number of commuters using the train or car from each municipality as the dependent variable and the distance between the municipality and the respective transport infrastructure as the independent variable. The coefficient obtained from the regression represents the effect of distance on commuters' mobility choices.

After obtaining the coefficient (statistically significant) from our Poisson regression analysis based on commuting data, we used it as a parameter within the impedance function. In particular, we maintained the general structure of the impedance function based on a negative exponential form of distance but replaced the coefficient with the significant value derived from the regression. This integration allowed us to account for the distance's effect on the probability of commuters' mode choice.

By using the Poisson regression coefficient within the impedance function, we were able to reconcile empirical results with Hansen's gravity model, thus ensuring a more solid foundation for calculating our accessibility indicator.

It is important to specify some caveats regarding the current case:

• The temporal distance considered is from the centroid of the municipality of residence to the nearest infrastructure, which may not necessarily be the one used by the commuter.

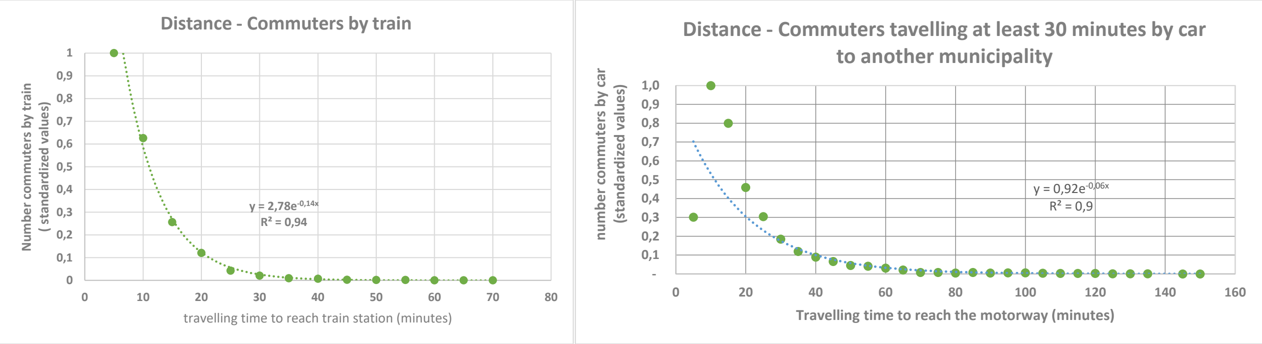
• The data refer to the commuting matrix of 2011, while the infrastructures are those reported for the years 2021 and 2022.

Additionally, the duration of the journey that the commuter will undertake is not taken into account. This choice stems from the fact that we are measuring the ability of that territory to attract a flow towards the railway infrastructure rather than the commuter's modal choice.

The estimation conducted (see Figure 11) found a coefficient (lambda) of 0.14 for railway stations and 0.06 for highway access points.

This result seems to support what has emerged from the literature and appears to endorse the choice of using a negative exponential function with a parameter (lambda) of 0.14 for railway stations and 0.06 for highway access points in the calculation of our accessibility measure.

FIGURE 11. RELATIONSHIP BETWEEN DISTANCE AND NUMBER OF COMMUTERS.



However, for airport and seaport infrastructures, there are no empirical data available to support a function or parameter. Therefore, given the lack of real data for precise estimates, it is appropriate to establish the decay function through a series of interventions based on "common sense" but also on the distribution and role of the infrastructures themselves under examination, in order to rebalance the discouragement effect in choosing a particular infrastructure due to distance and on-site presence.

This need can be met in two ways: by imposing a lambda coefficient closer to those found in the literature for other countries and therefore closer to a curve similar to that observed previously for railway stations (thus equal to 0.1), which is usually used when data are not available (Palacios and Le Geneidy 2022). Alternatively, one could also consider replacing the proposed negative exponential function with a simpler Gaussian function that has less impact for shorter distances and increases its impact as distance grows (see previous graph). In this way, the willingness of populations in municipalities located ten minutes away from the port or airport infrastructure will decrease by only 10%, at twenty-five to thirty minutes it will lose about 50% of the value of the services offered by the infrastructure, at 40 minutes it will lose 80%, and after 60 minutes it will lose 98%.

The results obtained in both cases still highlight a high correlation between the two different choices, which led us to choose a solution that involves the use of a Gaussian function.

For technical and methodological clarifications

Massimo Armenise
Massimo.Armenise@istat.it

Raffaella Chiocchini
Raffaella.Chiocchini@istat.it

Marianna Mantuano
Marianna.Mantuano@istat.it

Rossella Molinaro
Rossella.Molinaro@istat.it

Norina Salamone
Norina.Salamone@istat.it

Gianluigi Salvucci
Gianluigi.Salvucci@istat.it

Francesco Giovanni Truglia
Francescogiovanni.Truglia@istat.it