Networks and Spatial Economics
https://doi.org/10.1007/s11067-025-09674-2

RESEARCH

Hansen's Accessibility Theory and Machine Learning: a Potential Merger

Natalia Selini Hadjidimitriou[1] · Aura Reggiani[2] · John Östh[3] · Marco Mamei[1]

Accepted: 21 January 2025

Abstract
Accessibility is a central concept in transport geography, given its relationship with land development. It is defined as "the opportunity that an individual at a given location possesses to take part in a particular activity or set of activities" (Hansen, 1959). Hansen's Accessibility Model (HAM) can be computed using mobility flows between regions or employment data as a proxy for centres of attraction, coupled with an impedance function that incorporates travel costs. It has been the basis of multiple theoretical and empirical approaches over the years. In the last decades, advances in Machine Learning (ML) have also opened new possibilities for developing innovative approaches in the transport field. The objective of this work is primarily the study of the dynamics of urban accessibility, by considering two interrelated perspectives. Firstly, we aim to explore whether alternative data-based techniques, such as ML, can replicate the behaviour of HAM and thus capture, from data, the underlying theory linked to the spatial interaction model, by embedding the influence of geography, transport network, and socioeconomic factors on accessibility. Secondly, we investigate the feasibility of employing ML where flow data are unavailable, ensuring consistent measurements over time. A combined approach HAM-ML is developed and applied to this aim. As a case study, we examine inter-urban accessibility in two Italian regions, Lombardia and Emilia Romagna, based on socioeconomic and transport data from 2011. The results show the potential of this joint approach, opening new research prospects on accessibility from the theoretical, empirical, and policy viewpoints.

Keywords Accessibility · Hansen's model · Spatial interaction model · Machine learning · Random forest · Neural networks

Extended author information available on the last page of the article

Published online: 29 January 2025

1 Introduction

Accessibility allows reaching different destinations with the minimum effort and time. When the objective is to go to work, accessibility might impact employment opportunities and quality of life (Geurs et al. 2012; Litman 2013; Preston and Rajé 2007).

From the modelling viewpoint, we can consider Hansen (1959) as the ’founding father of accessibility’, thanks to his fundamental article, where he considers accessibility as: a) <*The potential of opportunities for interaction*>; b) <*A measure of the intensity of the possibility of interaction rather than just a measure of the ease of interaction*>; and c) <*accessibility is the measurement of the spatial distribution of activities about a point (area/city/region), adjusted for the ability and the desire of people or firms to overcome spatial separation*> (again Hansen 1959, p. 73).

Therefore, accessibility is strictly linked to spatial structures and the related social processes and behavioural rules.

Interestingly, from a theoretical viewpoint, Hansen’s Accessibility Model (HAM) can be linked to the Spatial Interaction Model (SIM), and thus to entropy maximization, as well as to connectivity networks (see Section 2).

Following Hansen, accessibility showed a growing popularity in research (see, e.g., the Swedish school, led by eminent scholars, such as Östh 2011; Mattsson and Weibull 1981; Weibull 1976), where accessibility can be seen as the ’hidden’ backbone of economic growth and development, linked to physical and virtual connectivity. In addition, the evolution of interaction and integration between spatial structures and human activities opens new research questions concerning the ’current’ role of accessibility, its analytical framework, and measurement.

Building on the above Hansen’s definition and interpretations of accessibility, we focus on the mobility flows measuring the spatial-economic opportunities (activities), given also the significant changes brought by the COVID-19 pandemic, i.e. the widespread adoption of teleworking. Relying solely on employment data may not accurately reflect accessibility patterns.

Measures relying on flow data, as often carried out (e.g., Kapatsila et al. 2023; Reggiani et al. 2011a), may consider the segments of ’virtual’ accessibility patterns. Therefore, we will implement the SIM for the computation of the HAM based on mobility flows among areas. However, when mobility flows’ data is unavailable, the adoption of an additional tool, such as Machine Learning (ML), would be useful to map out and forecast accessibility.

The underlying idea is that the increased computational capability and theoretical advances of ML might open new perspectives in HAM studies. While, in the past, some works used ML for forecasting spatial interaction and transport choices (Himanen et al. 1998; Mozolin et al. 2000; Reggiani and Nijkamp 1998), using ML to measure accessibility is still rare (Zhu and Diao 2020). Therefore, our work’s objective is to explore whether ML– if combined with the HAM– might provide new insights into accessibility, by capturing Hansen’s theory and forecasting accessibility patterns.

Consequently, the emerging theoretical issue is the following: Is ML able to learn from data to identify the underlying theoretical aspects of HAM? Or vice-versa: can ML provide a novel perspective on the HAM based on pattern recognition?

As an empirical platform for this novel approach, we will study the evolution of inter-urban accessibility in two regions of Northern Italy, i.e. Lombardia and Emilia Romagna. The idea is to model accessibility in 2011 (based on official data) to carry out temporal forecasts in 2020 (during the shock of COVID-19), when flow data are missing.

Starting from the above considerations, our adopted methodology is a combined approach HAM-ML. The first step is the construction of HAM based on the calibration of SIM, measured on the inter-urban mobility flows of the two Italian regions in 2011. This will be the target for the second methodological step, i.e. the development of an ML model to predict accessibility in 2020.

The work consists of the following sections: Section 2 introduces the theoretical background based on SIM which has a methodological link with HAM, and includes an overview of accessibility dimensions and interpretations. Section 2 also provides a literature review focusing on ML works in transport. Section 3 describes the methodology developed to execute the ML training and temporal experiments: accessibility performance in 2011 and related predictions in 2020, in the municipalities of the two mentioned Italian regions. Section 4 outlines the dataset and Section 5 illustrates the empirical analysis. Section 6 highlights the main conclusions in light of future research directions.

2 Methodological Background

This section offers an overview of the existing literature concerning accessibility, by highlighting the theoretical framework and models, as well as the fundamental accessibility measures and their intersection with machine learning (ML). Specifically, we describe HAM, which is grounded in the spatial interaction theory and, explore the works that have utilized ML in the transport sector.

2.1 Spatial Interaction Models

Spatial Interaction Models (SIMs) represent the interactions among different geographical areas and describe the relationship between an origin *i* and a destination *j*. In essence, SIMs assume that the flows T_{ij} between each origin and destination are influenced by the size of production and attraction areas and the impedance (distance d_{ij}, time t_{ij}, or cost c_{ij}) between them. Interestingly, Wilson's pioneering work (Wilson 1967, 1970) showed that SIM is the equilibrium solution of a consistent optimization problem: entropy maximization. This important demonstration provided a strong theoretical framework encompassing the analogy of SIMs with the simple gravity model (originally conceived for migration flows, see Ravenstein 1885). When the origin-destination (OD) matrix is known, SIMs predict more accurately the interactions and the model is called the production-attraction-constrained model or the doubly-constrained model (Fotheringham and O'Kelly 1989). The predicted interactions significantly improve when actual flows are known (Haynes and Fotheringham 2020; Wilson 1998). In particular, the doubly-constrained SIM is expressed and follows:

$$T_{ij} = A_i O_i B_j D_j f(t_{ij}) \tag{1}$$

where T_{ij} are the flows (trips), from origin i to destination j, O_i are the outflows (trips generated from origin i). D_j are the inflows (trips attracted to destination j), and $f(t_{ij})$ is the impedance function between i and j (here based on the travel time t_{ij}). A_i and B_j are balancing factors (for details, see Wilson 1967, 1970, 2021).
The impedance function $f(t_{ij})$ is often modeled using a negative-exponential or a negative-power function (see Section 2.2).

A comprehensive overview of various SIM formulations (unconstrained, production-constrained, attraction-constrained and doubly-constrained) can be found, among others, in Haynes and Fotheringham (2020); Fotheringham and O'Kelly (1989); Nijkamp and Reggiani (1992); Roy and Thill (2004).

Wilson noted that the balancing factor A_i emerging from a production-constrained SIM:

$$T_{ij} = A_i O_i D_j f(t_{ij}) \tag{2}$$

has the following specifications:

$$A_i = 1/\sum D_j f(t_{ij}) \tag{3}$$

A_i can be then interpreted as the inverse of accessibility since A_i reflects– in its denominator– HAM (see next Section 2.2). Since SIM Eq. 1 arises as an equilibrium solution from entropy maximization, Hansen's accessibility formulation can be linked to the system's entropy patterns (Hansen 1972).
A discussion on entropy as a 'universal' concept and theoretical construct that can unify the study of complex spatial economic systems and networks came to the fore last century and also recently (Reggiani et al. 2021). Accessibility linked to entropy highlights the fundamental role of entropy in spatial economic science, where accessibility can be recognized as a significant component, thanks to Eqs. 2 and 3.

All in all, these methodological considerations on SIMs and entropy provide theoretical frameworks for HAM, by paving the way for further insights and research. HAM will be illustrated in the next Section 2.2.

2.2 Hansen's Model of Accessibility

The Hansen's measure of accessibility is the first indicator that considers activities, land use and transportation network (Iacono et al. 2010). More specifically, according to Hansen (1959), the potential accessibility ACC_i of a zone i reflects the opportunities of all the other zones j, as follows:

$$ACC_i = \sum_j W_j t_{ij}^{-\gamma} \tag{4}$$

In formulation Eq. 4, ACC_i is the accessibility of zone *i* to the sum of opportunities W_j discounted by time-decay function $t_{ij}^{-\gamma}$. A destination can have social interaction opportunities W_j, such as activities of social practices (Järv et al. 2018), shopping opportunities (Guy 1983), employment and housing opportunities (Cervero et al. 1999), amenities (museums, theatres, etc.) or rural areas (Haynes and Fotheringham 2020). Therefore, accessibility has a social impact considering the relationship between transit and job accessibility as evidenced, among others, by Farber and Fu (2017). The time-decay function in Eq. 4 takes into account the attractiveness of a destination according to the negative-power form, as highlighted by Hansen (1959, p. 74). However, other types of the impedance function can be used, such as the negative-exponential function:

$$ACC_i = \sum_j W_j e^{-\beta t_{ij}} \tag{5}$$

According to Willigers et al. (2007), using a predefined shape function such as the power or the exponential is the best option. The choice of the impedance function will emerge from the calibration process (de Dios Ortúzar and Willumsen 2011). Reggiani et al. (2011a, 2011b) used five impedance function types for measuring inter-urban accessibility in German districts by showing that the power function seemed more suitable to model interaction in the case of hubs (i.e. Berlin, Hamburg, Koln, Munich), while the exponential function seemed adapted to a homogeneous area (Ruhr area). Interestingly, Hansen argues about the values of the exponent γ in Eq. 4, as follows: <*Unlike the contention of Stewart and others that this exponent should be unity, the assumption here is that the value of the exponent in this accessibility or potential model must be the same as that used in the gravity model*>. Therefore, we used a γ-value emerging from the calibration of SIM Eq. 1. In particular, from this calibration procedure, the negative-power form emerged as the best impedance function.

We will then consider an ML approach where the target model is HAM based on inflows D_j associated with a negative-power impedance function, comprehensive of the calibrated γ-value, as follows:

$$ACC_i = \sum_j D_j t_{ij}^{-\gamma} \tag{6}$$

The HAM-ML methodology will be illustrated in Section 3. The next section will focus on the multidimensional feature of accessibility.

2.3 Accessibility Dimensions

As previously mentioned, the first definition of accessibility has been provided by Hansen in 1959 in terms of the interaction between land use and transportation. However, there is a need to define accessibility based on time and space (Pirie 1979). An overview of accessibility measures (which can be reconducted to Hansen's formulation) highlights that multiple components of accessibility should be taken into account (Geurs and van Wee 2004). These can be classified into four main components of accessibility (transport, land-use, temporal and individual components) and four main perspectives (infrastructure-based, location-based, person-based, and utility-based). In addition, there is the need to interpret accessibility in terms of its socioeconomic dimension (Farrington and Farrington 2005; Geurs et al. 2012), as well as in terms of its relationship with the structure of the transport network. In this regard, the role of accessibility has been analyzed by considering two perspectives: the economy and the structural characteristics of the transport network. These two approaches generate consistent differences in terms of accessibility results, especially when different impedance functions are deployed (Reggiani et al. 2011a). This also underlines the importance of the distance decay function when measuring accessibility, especially when the size catchment area is large (Chen and Jia 2019).

Concerning the policy issues, the majority of transportation plans consider accessibility in terms of their goals such as better access to jobs or mobility which seemed to be the easiest to interpret (Boisjoly and El-Geneidy 2017; El-Geneidy et al. 2016; Geurs and van Wee 2004). In any case, the potential of accessibility indicators to guide the decision-making process can increase thanks to the use of GIS (Geographical Information Systems). These tools enable network-based accessibility measures which allow for the evaluation of the vulnerability of transportation (Chen et al. 2007).

From Table 1, we can see the multifaceted nature of accessibility. A comprehensive understanding of accessibility requires examining various dimensions, including

Table 1 Accessibility dimensions

Paper	Dimension	Accessibility
Farrington and Farrington (2005); Geurs et al. (2012)	Socio-economic	micro, macro and utlity-based
Boisjoly and El-Geneidy (2017); El-Geneidy et al. (2016); Geurs and van Wee (2004); Fan et al. (2012)	Policy	access to jobs or mobility
Chen et al. (2007); Farber et al. (2014); Handy and Clifton (2002); Kwan and Weber (2008)	Spatial	network-based, in the access to jobs or mobility at different scales (i.e., local, regional, or neighborhood levels)
Boisjoly and El-Geneidy (2017)	Temporal	dynamic accessibility

socioeconomic, policy, spatial, and temporal factors. In this work, we will consider all four dimensions of accessibility.

With reference to the scale measurement, accessibility can be measured at various levels, including individuals/groups, temporal and spatial scales. Recently, there is an increasing trend to measure accessibility at a disaggregate level (Geurs et al. 2012). Measurement of accessibility at the individual level involves segmenting the population (e.g., based on wages). For instance, Fan et al. (2012) showed that the creation of a new light rail service improved the accessibility of low wages employees who tend to utilize transit services more often than higher wages earners. Similarly, the temporal characteristics can be relevant for accessibility. For example, shops opening or working hours can influence accessibility patterns (Boisjoly and El-Geneidy 2017). In terms of spatial scale, measurement of accessibility can consider zone levels, such as census boundaries or transport planning zones (Farber et al. 2014).

Kwan and Weber (2008) proposed two main spatial scales of analysis: the regional and the local. From an accessibility perspective, the regional scale focuses on the accessibility to employment and the ease to reach these destinations. Conversely, at the local scale, the focus is on how the residential neighbourhoods should be redesigned to influence travel behaviour and encourage shorter walking trips. Indeed, among the factors that affect accessibility at neighbourhood level, there are activity patterns and the local transport system (Handy and Clifton 2002). However, there is a huge gap between data availability and accessibility factors at local level. In this regard, the ML approach could fill this gap mostly in the presence of small spatial areas.

2.4 Machine Learning in Transport: The Supervised Framework

ML has increased in popularity in the transport field in the last few years. In particular, ML has been deployed to estimate and predict transport efficiency, public transport costs, travel time, network flow, and so on.

Table 2 summarizes the various Machine Learning (ML) techniques applied to different transportation-related tasks. It highlights the diversity of ML methods, ranging from traditional ML algorithms like Random Forest (RF), Neural Networks (NN) and Support Vector Machines (SVM) to more advanced techniques like Reinforcement Learning (RL).

Overall, ML is widely applied to a variety of tasks such as estimating travel time, predict mode choice and optimizing urban mobility. Furthermore, this overview shows the growing interest in using ML to make decisions in the transport sector and highlights the importance of data availability needed to implement these techniques.

Given these successful ML applications in transport, in the present work, we deploy an ML approach, particularly a supervised ML to predict accessibility. The supervised framework consists of learning the relationship between the input and the target data on a sample dataset called a train set and making the prediction using an unseen dataset.

Examples of supervised learning algorithms are, among others, NN and RF. In particular, RF has been recently regarded as a highly effective method for making predictions in the transport sector, mostly in the mode choice (Cheng et al. 2019;

Table 2 ML in transportation: some examples of applications

Paper	ML applied to	ML method
Awad et al. (2023); Costa and Markellos (1997)	evaluate transport efficiency	NN
Ngo and Mishra (2023)	estimate travel time	NN
Liu et al. (2019)	estimate network flow	RF
Shafiq et al. (2020)	traffic classification	NN, RF
Chen and wan Zhang (2024)	urban transport optimization	RL
Heidari et al. (2022)	management of smart cities	NN
Ahmed and Diaz (2022)	urban mobility area (mode detection, localization based algorithms, and patterns recognition models)	RF, NN, SVM
Casali et al. (2022)	urban areas (land use, socio-economic factors, environment, infrastructure)	
Dougherty (1995)	driver behavior analysis, estimation of origin-destination matrices, pavement maintenance needs based on image analysis, and vehicle detection and classification	NN
Sekhar and Madhu (2016); Cheng et al. (2019)	mode choice predictions	RF
Wang et al. (2024)	travel demand	NN, RF, SVM

Sekhar and Madhu 2016). NN has a long history and wide applications in transport and spatial interaction/choice (Himanen et al. 1998; Nijkamp et al. 1996). Dougherty (1995) provides a comprehensive overview of NN applications in the transport sector, exploring their use in various domains, including driver behavior analysis, estimation of origin-destination matrices, pavement maintenance needs based on image analysis, and vehicle detection and classification (Table 2).

Given these premises, in the present work on accessibility theory and modelling, we adopt and compare two supervised ML techniques, such as RF and NN, which appear to be the most suitable ones. The ML details will be described in Section 3, while the next Section 2.5 will highlight the methodological insights emerging from a network accessibility approach and leading to ML adoption.

2.5 From Network Accessibility to Machine Learning: Some Considerations

The link between accessibility and network structures has been analyzed in the past years in light of the topological properties of a network. Accessibility embeds connectivity, as can be seen in formulations Eqs. 4 and 5 through the time/cost matrix t_{ij}, thus the network framework is an important feature of accessibility and vice-versa.

In this context, a significant role is played by the type of the impedance function. In particular, the power accessibility expressed by Eq. 4 and the exponential accessibility expressed by Eq. 5 seem to display different results. More precisely, experiments in the mobility network of the German districts showed a high correlation between the ranking of the most accessible areas implied by the power decay of Eq. 4 and the ranking of the connectivity indicators, such as the incoming connections and the betweenness. All in all, the hierarchy of accessibility Eq. 4 appeared to match the hierarchy of network connectivity by extrapolating the most important hubs for the first top districts (Reggiani et al. 2011a). However, this was not the case by adopting the exponential decay Eq. 5– or other decay forms– in the accessibility function, i.e. there was no matching between the hierarchical order of exponential accessibility vs. the network connectivity. In other words, the power accessibility Eq. 4 seemed to map spatial-economic disparities of the German commuting network modelled by a SIM, i.e. heterogeneity in the trip-makers and the mobility network. Exponential decay, which together with power decay constitutes the two most common forms of decay is better suited for decay in shorter, within-region trips (Östh et al. 2014, 2016). This result might have general theoretical implications, e.g. the compatibility between the related theoretical roots: spatial analysis (emerging from entropy theory) and network analysis (emerging from complexity theory), mostly in the presence of hub-network hierarchies.

A further consideration concerns the role of the parameter γ in Eqs. 4 and 5. It has been shown that this parameter γ is strictly related to the power law coefficient of the network as well as to the Rank Size rule coefficient (Reggiani et al. 2011a), and consequently interpreted as an index of concentration/agglomeration (Parr 1985). Unfortunately, in our analysis we cannot test the dynamics of γ, and thus the tendency– or not– towards a concentration process of our networks, having available only mobility data in 2011. We will then make use of ML to capture the dynamics of the network accessibility over the years.

ML has become increasingly used to study spatial and regional economic relationships (see for instance: Kopczewska 2022) but its use in the study of spatial accessibility is limited. In light of this, our theoretical/methodological research question is the following: is ML, based on HAM Eq. 4 as the target, and trained on HAM Eq. 6, able to capture the dynamics of γ, and, consequently, the dynamics of network accessibility?

Having said this, we will explore two ML techniques, RF and NN, as previously indicated. Their methodological characteristics will be briefly outlined in Section 3.

3 Machine Learning Vs. Hansen's Accessibility: The Adopted Methodology

Our case study is the accessibility analysis at the municipality level, in two Italian regions. Our idea is to train ML to rank municipalities based on their Hansen accessibility level. In other words, HAM Eq. 6 using inflows D_j, the travel time t_{ij} and the calibrated γ, is the baseline indicator. We have opted for this choice because we aim to assess accessibility in the context of the widespread adoption of remote working during the COVID-19 pandemic.

Initially, we calibrate a doubly-constrained SIM Eq. 1 utilizing 2011 inflows and travel time to determine the value of the time parameter γ to be inserted in HAM Eq. 6. The flows between each couple of municipalities and the travel time by car, together with the selection of the appropriate impedance function (in this case, the negative power) allow the values estimation of the γ parameter of the SIM. The flows start and end within the same region in a closed system. The estimation procedure, following the Newton-Raphson method, has been carried out through the accessibility component (SpinModel) software available at: https://resiliencetool.eu/SMHome.

Concerning the ML approach, the HAM measure Eq. 6, embedding inflows and the calibrated γ parameter in 2011, is the target variable/benchmark of ML. We used employment and travel time data from 2011 Eq. 4, by considering the value of $\gamma = 1$, as features for our ML model and trained it accordingly (for details, see Section 5.2).

We decided to investigate and compare two ML techniques: a) RF (Breiman 2001) developed in the 90s; b) NN, born in the 40s (McCulloch 1943), since these two ML techniques are the most consolidated. As Casali et al. (2022) highlight, NN and RF are the most popular supervised ML algorithms in urban research. The key difference between RF and NN lies in their learning approach. RF is an ensemble learning method that builds multiple decision trees from data and combines their outputs to make a final decision. In contrast, NN learns the complex, non-linear relationships within the data. In feedforward NN, information flows in one direction, from input to output, in the context of a training process that enables to learn the relationship between the input and the output (Fine 1999; Rosenblatt 1958; Scarselli and Tsoi 1998). The learning process is performed thanks to the backpropagation algorithm that calculates the gradient of a loss function (Werbos 2005). The feedforward NN consists of an input layer, one or more hidden layers, and an output layer, utilizing weights and activation functions to model complex patterns. The details about the architecture and hyperparameters of the adopted ML techniques are reported in Section 5.2.

We train both RF and NN to assess which of the two ML approaches better replicates Hansen's accessibility ranking. We trained each model ten times and considered the average. The evaluation of the RF and NN ability to predict accessibility considers three measures: the Spearman rank correlation, the Kendall rank correlation coefficient and the Normalized Discounted Cumulative Gain (NDCG).

Finally, we conduct a temporal experiment to assess the performance of the RF and NN in the year 2020 and to examine the influence of spatial-network and transport characteristics on the dynamics of accessibility. The next section details the dataset used in the experiments.

4 Dataset

The dataset consists of commuters' flows between municipalities located within two confining regions in the North of Italy: Lombardia and Emilia Romagna Fig. 1. It is interesting to know that the two regions have the common characteristic of having a

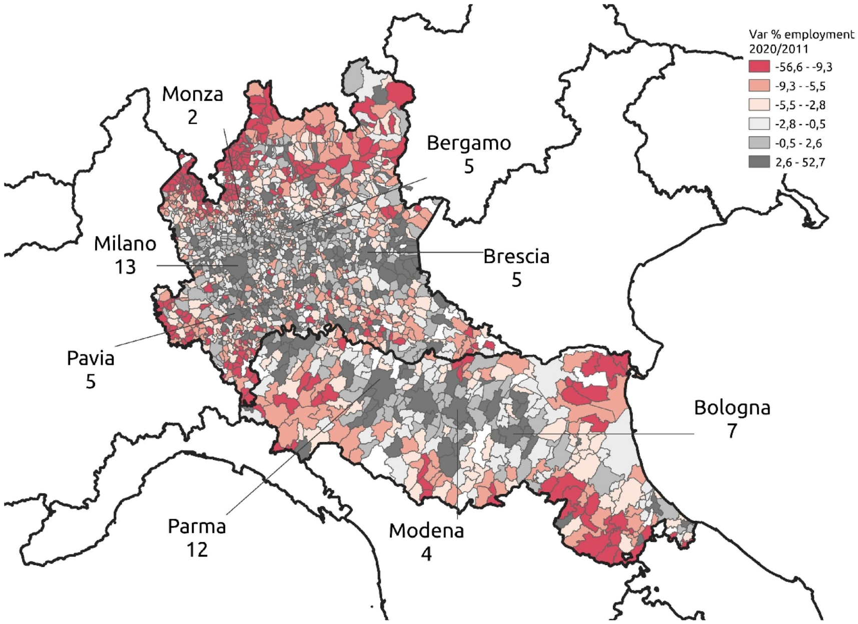


Fig. 1 Percentage variation in the number of employees in the two Italian Regions (2011–2020): Lombardia (on the top) and Emilia Romagna (on the bottom)

big metropolitan area that attracts commuters. According to ISTAT (Italian National Institute of Statistics)[1], the 31% of the GDP of all of Italy in 2020 was produced in Lombardia (22%) and Emilia Romagna (9%). In Lombardia, 9.7 million people live in 1,506 municipalities. Milan is the capital region, with 1.3 million inhabitants in 2011. In Emilia Romagna, there are 4.3 million inhabitants. Bologna is the capital region with 380,000 inhabitants and 333 municipalities in 2011.

The 15th Census of the population of 2011 by ISTAT (2011) was the source of information about commuter inflows and the number of employees. The considered inflows are trips performed by private cars (drivers)[2]. ISTAT measures the travel time using a commercially available software. Regarding the number of employees, we utilized, in addition to 2011, 2019 employment data (ISTAT 2019) and adjusted it based on regional statistics from 2020. Specifically, official statistics for the Lombardia and Emilia Romagna regions reported a 0.3% decrease in employment from 2019 to 2020[3] and we modified our employment data accordingly. As for travel time, we assumed the stability of car travel times over the nine-year period.

[1] www.istat.it

[2] The 2021 Origin-Destination Commuting Matrix was unavailable at the time this article was written. The updated matrix will only include data on the number of people commuting from an origin municipality to a destination municipality, omitting the additional variables published in 2011, such as travel time, mode of transport, and departure time. These omitted variables are scheduled for release in the first quarter of 2025, with details provided at the provincial level.

[3] statistica.regione.emilia-romagna.it, www.polis.lombardia.it,

Figure 1 displays the percentage variation in the number of employees in Lombardia and Emilia Romagna between 2011 and 2020. Remote areas, mainly those in the mountains, exhibit the most significant decrease in employment. Conversely, the two municipalities of Milan and Bologna (hubs/capital regions) demonstrate an increase in the number of employees from 2011 to 2020. An emerging question for future research is whether accessibility decreases in rural areas and increases in the hubs, after the COVID event.

We utilized the mentioned data to conduct the empirical analysis, which is detailed in the next section.

5 Empirical Analysis

This section outlines the empirical analysis, consisting of two main steps: i) calibration of the SIM to calculate HAM, as illustrated in Section 2.2; ii) execution of a temporal experiment, as described in Section 3. Each experiment consists of a comparative analysis between the HAM, the RF performance, and the NN performance regarding the hierarchical accessibility of the Italian municipalities in Lombardia and Emilia Romagna regions.

5.1 The Calibration of Spatial Interaction Models

Concerning step i), the calibration of a doubly-constrained SIM Eq. 1 related to the interurban transport flows in 2011, in the two Italian regions (Lombardia and Emilia Romagna), provided the γ-exponent values, is displayed in Table 4 (Appendix 1). The method used for the calibration of the γ-parameter has been indicated in Section 3.

Table 4 (Appendix 1) presents γ values of 1.640 for Lombardia and 1.455 for Emilia Romagna, aligning with some recent works (De Montis et al. 2007; Reggiani et al. 2011a, b), which suggest a range of 0.5 to 3 for interurban flows. The decrease of the γ-exponent in the Emilia Romagna region means that here travel time becomes a less restrictive factor (Hansen 1959).

It should be noted that these emerging γ-values will be introduced in Eq. 6, which is the target function of our ML approach (Section 5.2).

5.2 The Adopted Machine Learning

Concerning point ii), the ML modeling, the experiments consist of training both RF and NN on a set of features and using as a target variable the HAM developed in step i). The adopted RF is set to 1000 estimators, meaning that the result is the average of 1000 decision trees. The NN is a feedforward network with one hidden layer. To optimize our model's performance, we conducted a hyperparameter tuning process. We evaluated different numbers of neurons, ranging from 100 to 1000 with a step size of

100, and compared the effectiveness of Sigmoid and ReLU activation functions. The hidden layer have a rectified linear unit activation function and the output layer has a linear function. The hidden layer has 800 neurons, the number of iterations is 400, and the training interrupts if the loss function does not improve after ten iterations. The training algorithm is the Adaptive Moment Estimation (Adam) (Kingma and Ba 2017) and the loss function is the mean squared error.

The learning and temporal experiment deploy identical features, hyperparameters (of RF and NN) and type of NN. The first feature is the number of employees (E_i) in municipality *i*.

The second feature is HAM Eq. 4, where $W_j = E_j$ and $\gamma = 1$.

$$Att_i = \sum_{j,i \neq j} \frac{E_j}{t_{ij}} \tag{7}$$

The variable Att_i Eq. 7 measures the attractiveness of destinations, according to HAM Eq. 4. It should be noted that the target is HAM Eq. 6 in 2011, based on inflows and the calibrated γ values Table 4. In other words, the system, starting from E_i and Eq. 7, learns the accessibility and γ characteristics of Eq. 6.
Thus, our HAM-ML approach reads– in the training phase– as follows:

$$ACC_{2011} = ML(E_{2011}, Att_{2011}) \tag{8}$$

In Eq. 8, ACC_{2011} represents the ML model where HAM Eq. 6 is the target, measured using inflows and travel time in 2011, jointly with the calibrated γ values, while E_{2011} and Att_{2011} are the input data, i.e. E_i and Eq. 7 measured with 2011 data. Concerning the forecasting phase in 2020, our HAM-ML results to be as follows:

$$ACC_{2020} = ML(E_{2020}, Att_{2020}) \tag{9}$$

In Eq. 9, the ML predicts accessibility in 2020 (ACC_{2020}) based on E_i and Att_{2020} with employment updated to 2020.

5.3 Comparative Analysis

In this section, we illustrate the estimation of ML models' performance in learning to measure the HAM in 2011 (based on inflow data). Two features (E_i and Eq. 7) have been adopted to train the ML models.

The first three columns of Table 5 (Appendix 2) show the rank of the top fifteen and the ten bottom municipalities within the Lombardia region, in 2011. In general, the RF ranking of municipalities in Lombardia in the top position is more similar to the HAM ranking compared to NN. Concerning the bottom positions, RF and NN ranks differ from HAM, although RF does slightly better.

Moreover, the columns of Table 6 (Appendix 2) present the comparative analysis of the Emilia Romagna region in 2011. Here both RF and NN perform well in reproducing the HAM ranking in the top seven positions, while RF produces rankings similar to HAM in the bottom positions. However, since the latter five centres are small villages (around 800–1000 inhabitants) located in the mountains, accessibility is indeed low.

Due to missing data arising from the feature calculation process, the accessibility prediction is based on 1181 municipalities– instead of 1506 in Lombardia– and 305 municipalities in Emilia Romagna– instead of 330.

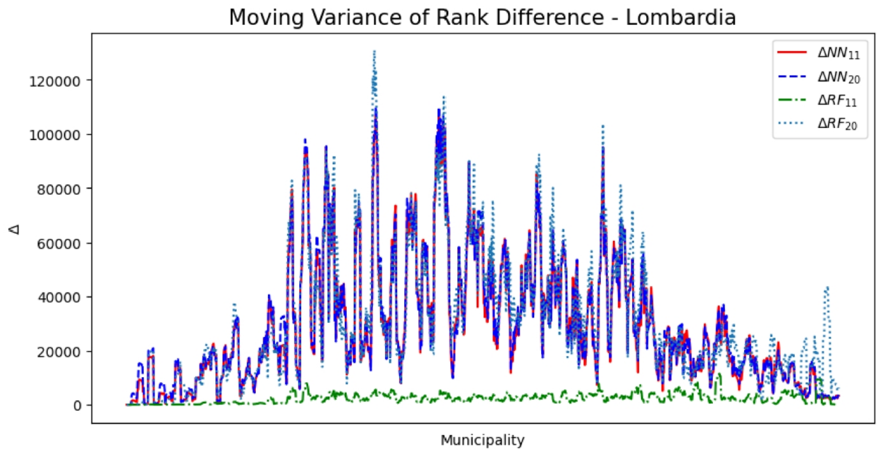
Subsequently, we consider three indicators reported in Table 3 to compare the rankings. While the literature usually considers Spearman, we measure two additional indicators to validate the results. The Spearman correlation (Zwillinger and Kokoska 1999) coefficient quantifies the strength and direction of association between two ranked variables. More specifically, it quantifies the degree of the linear relationship between the rank of one variable and the rank of another variable. We deploy this measure because the ranks are ordinal. The Spearman rank correlation is high when the relative positions of the observations in the rank are similar. It is a measure of the monotonicity of the relationship between two variables. Thus, if two ranks are identical, the value of ρ is one, indicating perfect monotonicity of the relationship. The Kendall Tau (Knight 1966) measures the number of concordant and discordant pairs between two rankings. Specifically, it assesses the correspondence between two rankings, where higher values indicate stronger agreement. Finally, the Normalized Discounted Cumulative Gain (NDCG) applies a logarithmic discount to the sum of scores ordered according to the predicted scores (Wang et al. 2013). The indicator ranges between 0 and 1 and returns high values when the first positions of the two rankings are similar. The NDCG is a measure of information retrieval. Therefore, the measure assigns greater importance to the similarity of higher rankings. In other words, the more similar the highest-ranked observations are, the higher is the value of the indicator.

As previously observed in Tables 5 and 6, the rankings most similar to the HAM are those produced by the RF, in the high as well as in the bottom positions.

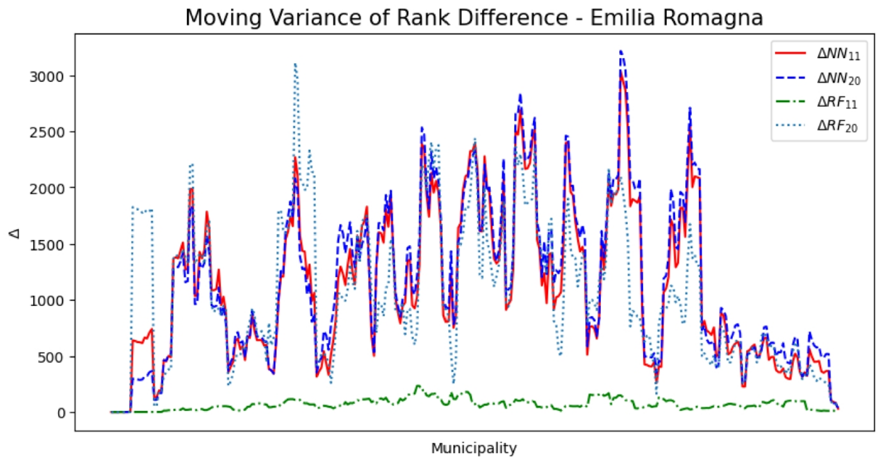
However, to understand if ranks are distributed similarly throughout the range of rank estimations we have to resort to other means of measure. If we sort accessibility according to the Hansen accessibility rank from greatest accessibility to lowest accessibility and subtract ranks of both RF and NN for the corresponding municipalities, we get results that can be used to study the variance of ranks from high to low

Table 3 Comparison between rankings

Region	ML	Spearman	Kendal	NDCG
Lombardia	NN	0.87	0.69	0.96
	RF	1.00	1.00	1.00
Emilia R.	NN	0.96	0.83	0.99
	RF	0.99	0.94	1.00



(a) Lombardia



(b) Emilia Romagna

Fig. 2 Moving variance (n=9) of rank difference between HAM and RF 2011 (ΔRF_{11}), RF2020 (ΔRF_{20}), NN 2011 (ΔNN_{11}) and NN 2020 (ΔNN_{20})

ranking municipalities. Figure 2 shows how a moving variance of rank difference develops throughout the rankings. The extremes represent the highest and lowest positions in the ranking. ΔRF_{11} (or ΔNN_{11}) denotes the moving variance, calculated over a 9-point window (n=9), of the difference between Hansen's ranking and the RF (or NN) ranking. Similarly, ΔRF_{20} (ΔNN_{20}) represents the moving variance (n=9) of the difference between the predicted 2020 ranking and Hansen's ranking. The result confirms that in particular RF rank difference (ΔRF_{11}) renders the least variance but that the rank difference variance increases slightly in the mid-ranks. For NN the patterns are similar but of much greater magnitude.

In essence, the comparative analysis highlights the power of ML as a forecasting tool. This ML ability was rather questioned in the past, in particular concerning NN trained to model spatial interaction with an exponential decay function (see, eg., Thill and Mozolin 2000). However, in our case study, where we employed RF to assess Hansen's accessibility (embedding a power-decay form), we can feel confident in affirming the predictive accuracy of ML, at least in the higher and bottom positions.

Based on these findings, we opted to deploy RF for our Temporal Experiments in 2020.

5.4 Temporal Experiment and Concluding Remarks

In our *Temporal Experiment*, we trained the RF model using the features described in Section 5.2, measured in 2011 within Lombardia and subsequently Emilia Romagna. We hypothesize that the ML models trained on 2011 data should effectively capture and reproduce HAM Eq. 6 under varying conditions. Consequently, they should be

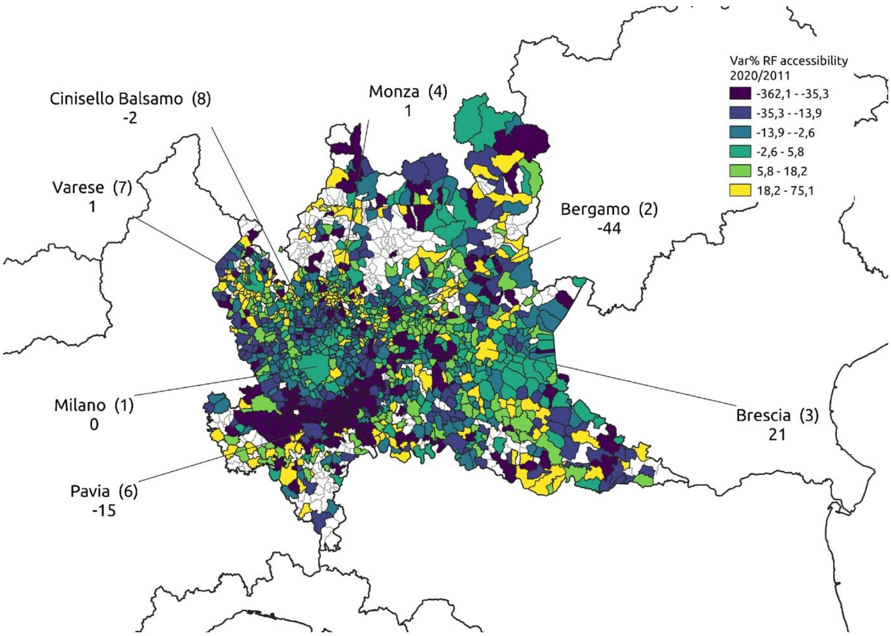


Fig. 3 Temporal experiment in the region of Lombardia: the map illustrates the percentage variation between 2011 and 2020 of predicted RF accessibility. The hierarchical order of RF accessibility of 2020 according to Table 5 is written in parentheses

able to predict the evolution of HAM across municipalities in both Lombardia and Emilia Romagna over time.

Figure 3 illustrates the percentage variation in accessibility between 2011 and 2020 in Lombardia. Overall, in the capital and satellite municipalities, the accessibility did not change much, such as in Milan (0%) or Monza (1%). Conversely, in the other municipalities of Regione Lombardia, a bit far from Milano, RF accessibility decreased, sometimes substantially, such as in Bergamo (−44%) or Pavia (−15%).

Concerning the hierarchy of the most accessible municipalities, Table 5 (Appendix 2) shows that the RF model of 2020 confirms, in general, Hansen's hierarchy of 2011, while NN displays a different hierarchical pattern. In particular, the hierarchical accessibility of the top urban areas seems to be rather stable over the years. Given the good performance of RF vs. NN in 2011, we are confident in the RF results of 2020.

Furthermore, we can observe that RF and HAM highlight in the top positions– in addition to the main cities of the Lombardia region– some satellite centres surrounding Milan. It should be noted that the main hubs around Milano are displayed according to a star network, thus, during the COVID time, the satellite centres around Milano have been preferred in accessibility since they are closer to Milano.

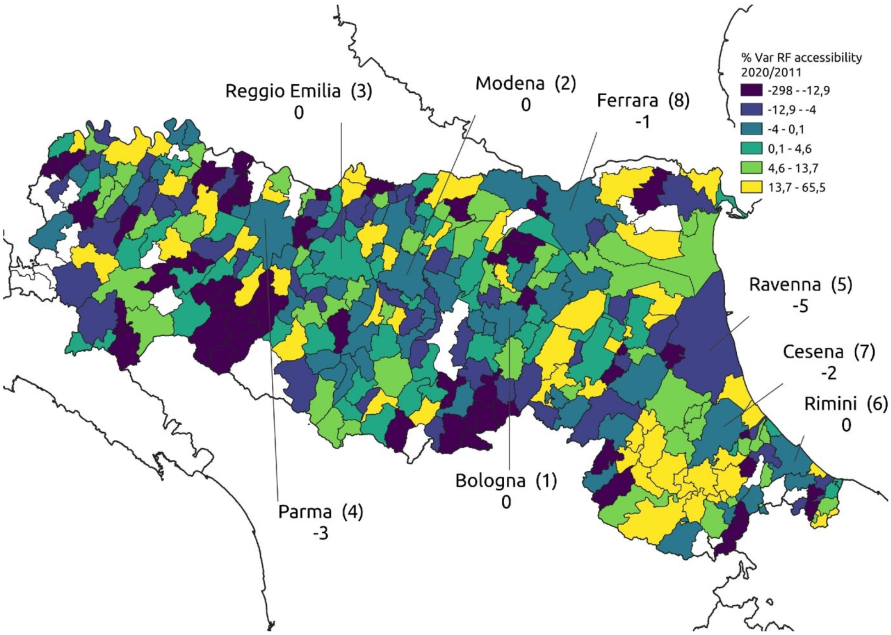


Fig. 4 Temporal experiment in the Region of Emilia Romagna: the map shows the percentage variation between 2011 and 2020 of predicted RF accessibility. The hierarchical order of RF accessibility of 2020 according to Table 6 is written in parentheses

Hubs are very relevant in a network structure, given their leading socio-economic characteristics, but they might also be critical: in the presence of shocks in the hubs, hubs might lead to the vulnerability or even disruption of the multiple connected links; however, hubs might create resilience in the restoration phase, thanks to their high accessibility (Reggiani 2022). Consequently, the right identification of hubs' hierarchy in network accessibility plays a fundamental role for planners and decision-makers.

Concerning the accessibility dynamics in the Emilia Romagna region, Table 6 shows the various rankings in 2020. Again RF, and not NN, matches Hansen's accessibility od 2011 in the top positions. It should be noted that the Emilia Romagna Region displays a linear topological structure in its hubs. Also in this case, the four main hubs show stable accessibility over the years.

Figure 4 shows the RF variation in accessibility between 2011 and 2020 in Emilia Romagna. Analogously to the Lombardia region, accessibility remained stable in the main centres, such as in Bologna (0%), Modena (0%) or Reggio Emilia (0%). The stability of accessibility in the satellite areas during the COVID pandemic is in line with the employment stability observed in the same areas, as shown in Fig. 1. In other words, in Regione Emilia Romagna, the main cities maintained their hubs' role in accessibility, given their good (linear) infrastructure connectivity.

Finally, it is interesting to consider the changes in the labour market, particularly with the current increase in remote working. As previously pointed out, in the two considered regions, employment decreased by about 0.3% in 2020 (with respect to 2019), although occupation has increased in the main and satellite centres compared to 2011 (Fig. 1; ISTAT 2011).

If we are subjecting the 2020 temporal experiment datasets to a moving variance of rank difference similar to the methodology illustrated in Section 5.3, it can be seen that the related RF results become remarkably different in the mid positions from the corresponding 2011 version Fig. 2. In general, for all the series (both 2011 and 2020) the variance increases in the mid-ranks suggesting that top and low accessibility remains stable regardless of the method used for accessibility estimation, while the mid-sections are more diverse from the Hansen's ranks. Since the 2020 ML models predict accessibility during the pandemic, where we have no flow data, we should be cautious with the emerging RF hierarchies in the mid-accessibility positions. Consequently, we might consider these RF findings jointly with Hansen's outcomes as a band of results to be analyzed, in light of policy actions on network accessibility improvements.

6 Conclusion and Future Research

The objective of this study was twofold. Firstly, we aimed to explore whether alternative data-based techniques, such as Machine Learning (ML), could extrapolate and replicate the behavior of Hansen Accessibility Model (HAM) and thus capture, from data, the underlying theory linked to the spatial interaction model, by embedding the influence of geography, transport network, and socioeconomic factors on accessibility. Secondly, we experimented the feasibility of employing ML in situations where flow data are unavailable.

Among ML techniques, we selected Random Forests (RF) and Neural Networks (NN) due to their methodological characteristics, such as the possibility to tune hyperparameters. As a case study, we examined accessibility at the municipality level in two Italian regions, Lombardia and Emilia Romagna, based on employment and (inter-urban) transport data from the year 2011. In this context, we considered HAM using inflows and travel time as the target variable and conducted a temporal experiment in the year 2020, aiming to capture the dynamics of network structures (hierarchy of hubs) underlying accessibility.

Through the mentioned temporal experiment in 2020, we predicted RF and NN accessibility in scenarios where flow data were unavailable. The comparative analysis revealed that RF appeared to reproduce Hansen's accessibility ranking in 2011 better– and thus to grasp a fundamental component of the related entropy pattern. In particular, the emerging RF results of Lombardia and Emilia Romagna regions show that accessibility towards the main hubs remains stable– in its hierarchical order– during the COVID time (the year 2020), while NN captures a different hierarchical order of the municipalities. Given the good performance indicators of RF vs. NN in 2011, the 2020 RF hierarchy of hubs seems plausible.

Hubs are very relevant in a network structure, given their leading socio-economic role, but also critical in the presence of shocks: because of their multiple links with the other nodes of the networks, hubs might cause the vulnerability/disruption of the whole network; however, hubs can also be very resilient, thanks to accessibility in the restoration phase. Consequently, the right identification of hub hierarchy in a network plays a fundamental role for planners and decision-makers.

It should be noted that in the Lombardia region the main hubs around Milano are displayed according to a star network, thus, during the COVID time, the satellite centres around Milano are also preferred in accessibility since they are closer to Milano, and emerge in the first ten-top positions. Analogously, in the Emilia Romagna region, RF in 2020 seems to match the ranking of 2011, i.e. the main socioeconomic and mobility hubs. Interestingly, the Emilia Romagna Region displays a linear connectivity structure in its hubs. Also in this case, in 2020, we can see stability in the accessibility ranking of the hubs.

These findings and considerations pave the way for in-depth exploration, also from the theoretical viewpoint, of accessibility in relation to the various topologies underlying the socio-economic mobility networks. All in all, the combined approach of HAM-ML, by providing a band of results, especially for the mid-accessibility positions, can be considered a powerful tool for policy analysis on network accessibility and is worth further investigation. In this regard, based on the performed RF/NN training in the Lombardia and Emilia-Romagna regions, future research consists of testing spatial experiments, i.e. to propose and implement the HAM-ML approach in other Italian regions, as well as in different European regions, which might show similar geographical patterns, e.g. high spatial-economic differences between the capital region and peripheral centres. This HAM-ML approach can also be conducted in further temporal dimensions to measure the post-COVID hierarchical accessibility at both intra- and inter-urban levels.

Integration of the HAM-ML accessibility tool with other approaches, such as complex network analysis and resilience/vulnerability, are finally considered for future research. Theoretically, the relationship between network connectivity and accessibility is not well established. Complex network analysis can provide a series of network connectivity indicators (starting from the well-known random network and scale-free network indicators). A formal correspondence between connectivity indicators and accessibility indicators would then be interesting, in parallel with the empirical HAM-ML analysis. In other words, we might conjecture that ML can capture not only Hansen's accessibility but also the associated entropy pattern, as well as the topological connectivity underlying the system under analysis. The same holds for resilience/vulnerability analysis and accessibility. Recent experiments in this regard show that the main hubs (e.g. the capital cities) are the most accessible and the most resilient, as previously highlighted.

Consequently, theoretical and empirical studies on the dynamics of the hierarchical accessibility and connectivity of the centers in a certain system/region might be relevant to transportation planning strategies, in light of resilience objectives. The HAM-ML approach can provide useful insights into the challenges of this framework.

Appendix 1 Calibration of SIM

Table 4 Time (γ) parameter values emerging from the calibration of SIM - year 2011

Italian regions	γ
Lombardia	1.640
Emilia Romagna	1.455

Appendix 2 HAM-ML Accessibility Experiments

In this annex, we report a comparative analysis of Hansen's accessibility, RF and NN performance in 2011 and 2020, in both Lombardia and Emilia Romagna regions (Tables 5 and 6).

Table 5 Accessibility models in the Lombardia municipalities (Italy, 2011 & 2020)

Rank	HAM 2011 (Inflows)	RF Model 2011	NN Model 2011	RF Model 2020	NN Model 2020
1	**Milano**	**Milano**	**Monza**	**Milano**	**Monza**
2	**Brescia**	**Brescia**	Sesto San Giovanni	**Bergamo**	Sesto San Giovanni
3	**Bergamo**	**Monza**	**Milano**	**Brescia**	**Milano**
4	**Monza**	**Bergamo**	Paderno Dugnano	**Monza**	Paderno Dugnano
5	Varese	Sesto San Giovanni	Cinisello Balsamo	Sesto San Giovanni	Cinisello Balsamo
6	Cinisello Balsamo	Varese	Bresso	Pavia	Bresso
7	Sesto San Giovanni	Cinisello Balsamo	Cologno Monzese	Varese	Cologno Monzese
8	Pavia	Paderno Dugnano	Cormano	Cinisello Balsamo	Cormano
9	Busto Arsizio	Pavia	Nova Milanese	Paderno Dugnano	Nova Milanese
10	Como	Como	Novate Milanese	Cesano Maderno	Novate Milanese
11	Legnano	Rho	Brugherio	Desio	Brugherio
12	Gallarate	Legnano	Desio	Limbiate	Desio
13	Segrate	Busto Arsizio	Muggiò	Cremona	Muggiò
14	Rho	Desio	Vimodrone	Como	Vimodrone
15	Agrate Brianza	Cologno Monzese	Cusano Milanino	Lodi	Cusano Milanino
...	...	...	...	...	...
1172	Dumenza	Lanzada	Genivolta	Campione d'Italia	Genivolta
1173	Quingentole	Caspoggio	Viadanica	Valbondione	Viadanica
1174	Magnacavallo	Monno	Quingentole	Ponte Nizza	Quingentole
1175	Lanzada	Valbondione	Zone	Esino Lario	Zone
1176	Torre di Santa Maria	Scandolara Ravara	Valbondione	Limone sul Garda	Valbondione
1177	Caspoggio	Tignale	Campodolcino	Candia Lomellina	Campodolcino
1178	Villa di Chiavenna	Zavattarello	Oltre il Colle	Sartirana Lomellina	Oltre il Colle
1179	Monno	Esino Lario	Zavattarello	Oltre il Colle	Zavattarello
1180	Campodolcino	Limone sul Garda	Ponte Nizza	Quingentole	Ponte Nizza
1181	Campione d'Italia	Campodolcino	Esino Lario	Zone	Esino Lario

Table 6 Temporal experiments of accessibility in the Emilia Romagna municipalities (Italy, 2011 & 2020)

Rank	HAM 2011 (Inflows)	RF Model 2011	NN Model 2011	RF Model 2020	NN Model 2020
1	**Bologna**	**Bologna**	**Bologna**	**Bologna**	**Bologna**
2	**Modena**	**Modena**	**Modena**	**Modena**	**Modena**
3	**Reggio Emilia**	**Reggio Emilia**	San Lazzaro di Savena	**Reggio Emilia**	San Lazzaro di Savena
4	**Parma**	**Parma**	Casalecchio di Reno	**Parma**	Casalecchio di Reno
5	Ravenna	Ravenna	**Reggio Emilia**	Ravenna	**Reggio Emilia**
6	Forlì	Forlì	Formigine	Rimini	Formigine
7	Rimini	Rimini	Ravenna	Cesena	Castelfranco Emilia
8	Cesena	Ferrara	Castelfranco Emilia	Ferrara	Ravenna
9	Ferrara	Cesena	**Parma**	Forlì	**Parma**
10	Piacenza	Imola	Bagnacavallo	Imola	Castenaso
11	Imola	Carpi	Castenaso	Carpi	Ferrara
12	Carpi	Piacenza	Ferrara	Faenza	Bagnacavallo
13	Faenza	Faenza	Cesena	Formigine	Cesena
14	Sassuolo	Formigine	Faenza	Reggiolo	San Giovanni in Persiceto
15	Casalecchio di Reno	Sassuolo	Forlì	Casalecchio di Reno	Faenza
...	...	...	...	...	...
296	Compiano	Riolunato	Varsi	Premilcuore	Varsi
297	Farini	Varsi	Montegridolfo	Riolunato	Ferriere
298	Riolunato	Palanzano	Ferriere	Mondaino	Montegridolfo
299	Morfasso	Montegridolfo	Palanzano	Monchio delle Corti	Farini
300	Premilcuore	Farini	Pellegrino Parmense	Montegridolfo	Pellegrino Parmense
301	Palanzano	Premilcuore	Tornolo	Farini	Tornolo
302	Fiumalbo	Morfasso	Morfasso	Ferriere	Morfasso
303	Monchio	Tornolo	Monchio delle Corti	Morfasso	Monchio delle Corti
304	Tornolo	Monchio delle Corti	Premilcuore	Pellegrino Parmense	Portico e San Benedetto
305	Ferriere	Ferriere	Portico e San Benedetto	Tornolo	Premilcuore

Acknowledgements The authors would like to thank two reviewers for their valuable comments and insightful feedback.

Author Contributions N.S.H. and A.R. wrote the main manuscript text and made the experiments; N.S.H. made Figs. 1, 3 and 4; M.M. and J.O. contributed to the experiments; J.O. made Fig. 2 and contributed to the comparative analysis. All authors reviewed the manuscript.

Funding Open access funding provided by Università degli Studi di Modena e Reggio Emilia within the CRUI-CARE Agreement.
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability The dataset used in this work is available at the following link: https://figshare.com/s/dbdf5fc2ae662644d714.

Declarations

Competing interests The authors declare no competing interests.

References

Ahmed DB, Diaz EM (2022) Survey of machine learning methods applied to urban mobility. IEEE Access 10:30349–30366

Awad FA, Graham DJ, AitBihiOuali L, et al (2023) Benchmarking the performance of urban rail transit systems: a machine learning application. Transportmetrica A: Transport Science pp 2241566

Boisjoly G, El-Geneidy A (2017) How to get there? a critical assessment of accessibility objectives and indicators in metropolitan transportation plans. Transport Policy 55:38–50

Breiman L (2001) Random forests. Mach Learn 45(1):5–32

Casali Y, Aydin NY, Comes T (2022) Machine learning for spatial analyses in urban areas: a scoping review. Sustainable Cities Soc 85:104050

Cervero R, Rood T, Appleyard B (1999) Tracking accessibility: Employment and housing opportunities in the san francisco bay area. Environ Plan A: Economy Space 31(7):1259–1278

Chen A, Yang C, Kongsomsaksakul S et al (2007) Network-based accessibility measures for vulnerability analysis of degradable transportation networks. Netw Spatial Econ 7(3):241–256

Chen G, wan Zhang J (2024) Intelligent transportation systems: Machine learning approaches for urban mobility in smart cities. Sustain Cities Soc 107:105369

Chen X, Jia P (2019) A comparative analysis of accessibility measures by the two-step floating catchment area (2sfca) method. Int J Geographical Inf Sci 33(9):1739–1758

Cheng L, Chen X, De Vos J et al (2019) Applying a random forest method approach to model travel mode choice behavior. Travel Behaviour Soc 14:1–10

Costa Á, Markellos RN (1997) Evaluating public transport efficiency with neural network models. Trans Res Part C: Emerging Technol 5(5):301–312

De Montis A, Barthélemy M, Chessa A et al (2007) The structure of interurban traffic: a weighted network analysis. Environ Plan B: Plan Design 34(5):905–924

de Dios Ortúzar J, Willumsen L (2011) Modelling transport. John wiley & sons

Dougherty M (1995) A review of neural networks applied to transport. Trans Res Part C: Emerging Technol 3(4):247–260

El-Geneidy A, Levinson D, Diab E et al (2016) The cost of equity: Assessing transit accessibility and social disparity using total travel cost. Trans Res Part A: Policy Practice 91:302–316

Fan Y, Guthrie A, Levinson D (2012) Impact of light rail implementation on labor market accessibility: A transportation equity perspective. J Trans Land Use 5(3)

Farber S, Fu L (2017) Dynamic public transit accessibility using travel time cubes: Comparing the effects of infrastructure (dis)investments over time. Comput Environ Urban Syst 62:30–40

Farber S, Morang M, Widener M (2014) Temporal variability in transit-based accessibility to supermarkets. Appl Geography 53:149–159

Farrington J, Farrington C (2005) Rural accessibility, social inclusion and social justice: towards conceptualisation. J Transport Geography 13(1):1–12. Sustainability and the Interaction Between External Effects of Transport (Part Special Issue, pp 23–99)

Fine TL (1999) Feedforward neural network methodology. Springer Science & Business Media

Fotheringham A, O'Kelly M (1989) Spatial interaction models: formulations and applications, vol 1. Kluwer Academic Publishers Dordrecht

Geurs K, van Wee B (2004) Accessibility evaluation of land-use and transport strategies: review and research directions. J Trans Geography 12(2):127–140

Geurs K, Krizek K, Reggiani A (2012) Accessibility analysis and transport planning: an introduction. Accessibility Analysis and Transport Planning. Edward Elgar Publishing, Cheltenham, UK, pp 1–12

Guy C (1983) The Assessment of Access to Local Shopping Opportunities: A Comparison of Accessibility Measures. Environ Plan B 10(2):219–237

Handy S, Clifton K (2002) Evaluating neighborhood accessibility: Possibilities and practicalities. J Trans Stat 4(23):67–78

Hansen S (1972) Utility, accessibility and entropy in spatial modelling. The Swedish J Econ pp 35–44

Hansen W (1959) How accessibility shapes land use. J Am Institute Plan 25(2):73–76

Haynes K, Fotheringham A (2020) Gravity and spatial interaction models. In: Web Book of Regional Science. Regional Research Institute, West Virginia University

Heidari A, Navimipour NJ, Unal M (2022) Applications of ml/dl in the management of smart cities and societies based on new trends in information technologies: A systematic literature review. Sustain Cities Soc 85:104089

Himanen V, Nijkamp P, Reggiani A (1998) Neural networks in transport applications. Ashgate Pub Ltd

Iacono M, Krizek K, El-Geneidy A (2010) Measuring non-motorized accessibility: issues, alternatives, and execution. J Transport Geography 18(1):133–140

ISTAT (2011) Census. Italy National Institute of Statistics

ISTAT (2019) Census. Italy National Institute of Statistics

Järv O, Tenkanen H, Salonen M et al (2018) Dynamic cities: Location-based accessibility modelling as a function of time. Appl Geography 95:101–110

Kapatsila B, Palacios MS, Grisé E et al (2023) Resolving the accessibility dilemma: Comparing cumulative and gravity-based measures of accessibility in eight canadian cities. J Transport Geography 107:103530

Kingma D, Ba J (2017) Adam: A method for stochastic optimization. arXiv:1412.6980

Knight W (1966) A computer method for calculating kendall's tau with ungrouped data. J Am Stat Assoc 61(314):436–439

Kopczewska K (2022) Spatial machine learning: new opportunities for regional science. Ann Regional Sci 68(3):713–755

Kwan M, Weber J (2008) Scale and accessibility: Implications for the analysis of land use-travel interaction. Appl Geography 28(2):110–123

Litman T (2013) The new transportation planning paradigm. Ite J-institute Transportation Eng 83

Liu Z, Liu Y, Meng Q et al (2019) A tailored machine learning approach for urban transport network flow estimation. Trans Res Part C: Emerging Technol 108:130–150

Mattsson L, Weibull J (1981) Competition and accessibility on a regional labour market. Regional Sci Urban Econ 11(4):471–497

McCulloch W (1943) A logical calculus of ideas imminent in nervous activity. The bulletin of mathematical biophysics

Mozolin M, Thill J, Usery E (2000) Trip distribution forecasting with multilayer perceptron neural networks: A critical evaluation. Trans Res Part B: Methodological 34(1):53–73
Ngo H, Mishra S (2023) Traffic graph convolutional network for dynamic urban travel speed estimation. Netw Spatial Econ 23(1):179–222
Nijkamp P, Reggiani A (1992) Interaction, Evolution and Chaos in Space. Springer
Nijkamp P, Reggiani A, Tritapepe T (1996) Modelling inter-urban transport flows in italy: A comparison between neural network analysis and logit analysis. Trans Res Part C: Emerging Technol 4(6):323–338
Östh J (2011) Introducing a method for the computation of doubly constrained accessibility models in larger datasets. Netw Spatial Econ 11(4):581–620
Östh J, Reggiani A, Galiazzo G (2014) Novel methods for the estimation of cost–distance decay in potential accessibility models. In: Accessibility and spatial interaction. Edward Elgar Publishing, p 15–37
Östh J, Lyhagen J, Reggiani A et al (2016) A new way of determining distance decay parameters in spatial interaction models with application to job accessibility analysis in sweden. European J Transport Infrastructure Res 16(2):344–363
Parr J (1985) A population-density approach to regional spatial structure. Urban Studies 22(4):289–303
Pirie G (1979) Measuring accessibility: A review and proposal. Environ Planning A: Econ Space 11(3):299–312
Preston J, Rajé F (2007) Accessibility, mobility and transport-related social exclusion. J Transport Geography 15(3):151–160
Ravenstein EG (1885) The laws of migration. J Stat Soc London 48(2):167–235
Reggiani A (2022) The architecture of connectivity: A key to network vulnerability, complexity and resilience. Netw Spatial Econ 22(3):415–437
Reggiani A, Nijkamp P (1998) The economics of complex spatial systems. Elsevier
Reggiani A, Bucci P, Russo G (2011) Accessibility and network structures in the german commuting. Netw Spatial Econ 11(4):621–641
Reggiani A, Bucci P, Russo G (2011) Accessibility and impedance forms: Empirical applications to the german commuting network. Int Regional Sci Rev 34(2):230–252
Reggiani A, Schintler LA, Czamanski D, et al (2021) Handbook on Entropy, Complexity and Spatial Dynamics: A Rebirth of Theory? Edward Elgar Publishing
Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review 65(6):386
Roy J, Thill J (2004) Spatial interaction modelling. In: Florax R, Plane D (eds) Fifty Years of Regional Science. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp 339–361
Scarselli F, Tsoi AC (1998) Universal approximation using feedforward neural networks: A survey of some existing methods, and some new results. Neural Netw 11(1):15–37
Sekhar C, Madhu E (2016) Mode choice analysis using random forrest decision trees. Transportation Research Procedia 17:644–652
Shafiq M, Tian Z, Bashir AK et al (2020) Data mining and machine learning methods for sustainable smart cities traffic classification: A survey. Sustain Cities Soc 60:102177
Thill J, Mozolin M (2000) Feedforward neural networks for spatial interaction: Are they trustworthy forecasting tools? In: Reggiani A (ed) Spatial economic science: new frontiers in theory and methodology. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp 355–381
Wang S, M B, Zheng Y et al (2024) Comparing hundreds of machine learning and discrete choice models for travel demand modeling: An empirical benchmark. Transportation Research Part B: Methodological 190:103061
Wang Y, Wang L, Li Y, et al (2013) A theoretical analysis of ndcg ranking measures. In: Proceedings of the 26th annual conference on learning theory (COLT 2013), p 6
Weibull J (1976) An axiomatic approach to the measurement of accessibility. Regional Sci Urban Econ 6(4):357–379
Werbos PJ (2005) Applications of advances in nonlinear sensitivity analysis. In: System Modeling and Optimization: Proceedings of the 10th IFIP Conference New York City, USA, August 31–September 4, 1981, Springer, pp 762–770
Willigers J, Floor H, van Wee B (2007) Accessibility indicators for location choices of offices: An application to the intraregional distributive effects of high-speed rail in the netherlands. Environ Plan A: Econ Space 39(9):2086–2898
Wilson A (1967) A statistical theory of spatial distribution models. Trans Res 1(3):253–269

Wilson A (1970) Entropy in urban and regional modelling. Pion, London; re-published by Routledge, Abingdon
Wilson A (1998) Land-use/transport interaction models: Past and future. J Transport Econ Policy 32(1):3–26
Wilson A (2021) Entropy in urban and regional modelling. In: Handbook on Entropy, Complexity and Spatial Dynamics. Edward Elgar Publishing, p 20–31
Zhu Y, Diao M (2020) Crowdsourcing-data-based dynamic measures of accessibility to business establishments and individual destination choices. Trans Res Part D: Transport Environ 87:102382
Zwillinger D, Kokoska S (1999) CRC standard probability and statistics tables and formulae. Crc Press

Authors and Affiliations

Natalia Selini Hadjidimitriou[1] · Aura Reggiani[2] · John Östh[3] · Marco Mamei[1]

✉ Natalia Selini Hadjidimitriou
selini@unimore.it

[1] Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Reggio Emilia, Italy

[2] Department of Economics, University of Bologna, Piazza Scaravilli, Bologna, Italy

[3] Department of Built Environment/TKD, Oslo Metropolitan University, Oslo, Norway