Revolutionizing Mobility: Big Data Applications in Transport Planning

ANTONELLA FALANGA[1,*], ARMANDO CARTENÌ[2]
[1]Engineering Department,
University of Campania "Luigi Vanvitelli",
via Roma 29, 81031 Aversa (Caserta),
ITALY

[2]Department of Architecture and Industrial Design,
University of Campania "Luigi Vanvitelli",
via S. Lorenzo 31, 81031 Aversa (Caserta),
ITALY

Corresponding Author

Abstract: - Today an unprecedented amount of data coming from several sources, including mobile devices, sensors, tracking systems, and online platforms, characterizes our lives. The term "big data" not only refers to the quantity of data but also to the variety and speed of data generation. These data hold valuable insights that, when extracted and analyzed, facilitate informed decision-making. The 4Vs of big data - velocity, volume, variety, and value - highlight essential aspects, showcasing the rapid generation, vast quantities, diverse sources, and potential value addition of this kind of data. Big data's pervasive impact enhances societal aspects, elevating the quality of life, service efficiency, and problem-solving capacities. However, during this transformative era, new challenges arise, including data quality, privacy, data security, cybersecurity, interoperability, the need for advanced infrastructures, and staff training. Within the transportation sector (the topic investigated in this research), applications span planning, designing, and managing systems and mobility services. Among the most common big data applications within the transport sector, there are, for example, real-time traffic monitoring, bus/freight vehicle route optimization, vehicle maintenance, road safety, and all the autonomous and connected vehicles applications, in addition to the travel demand estimation useful for a sustainable transportation planning. Emerging technologies, offering substantial big data at lower costs than traditional methods, play a pivotal role in this context. Starting from these considerations, the present study explores two recent Italian big-data applications within the transport sector starting from the database of the Italian Ministry of Infrastructure and Transport and the Ministry of Health. The first one investigates the proper national demand estimation by transport mode and territorial area of interest, while the second one correlates the diffusion of the COVID-19 pandemic with the mobility habits in the Country. The lessons learned from these case studies are: *i)* the large amount of mobility data is useful for estimating mobility habits as long as they are adequately treated (e.g. high professional skills are necessary) to certify the quality of the data; furthermore, also multi-source and multi-format data can significantly contribute to a better knowledge of the phenomenon, but only if they are adequately archived and processed; *ii)* the large amount of data made available to the different (many) operators/institutions has made possible to correlate the spread of the pandemic with the behavior of citizens; concerning transport sector, was observed that the daily mobility habits influence infections registered three weeks later and areas with higher transport accessibility are more rapidly affected by infections.

Key-Words: - Big data, cloud computing, database, decision-making, public transport, sustainable mobility, transport demand, transportation planning, Covid-19 pandemic, connected vehicles, autonomous vehicles.

Received: March 27, 2023. Revised: October 23, 2023. Accepted: December 17, 2023. Published: December 31, 2023.

1 Introduction

In recent years, there has been an extensive availability of data sourced from entities, sensors, portable gadgets, the Internet of Things (IoT), multimedia, social networks, business transactions, and scientific data, [1], [2]. This copious volume of data, known as big data, refers to extremely large and complex data sets that require specific techniques to be efficiently collected, managed, processed, and analyzed, [3], [4]. Moreover, they possess promising market value and have recently captured substantial attention from scientific circles, governmental bodies, and industries owing to their significantly positive influence across various domains. This surge of information has revolutionized many sectors, such as business for improving decision-making processes, healthcare for clinical record analysis and medical research, education for enhancing teaching methodologies, agriculture for optimizing crop management, finance for risk assessment and fraud detection, media and entertainment for personalized content recommendations, emergency for a real-time response during crisis/events, and also mobility for the urban planning and for the design/management of public and private transport services.

From the literature review emerges that there are three main characteristics associated with big data, known as the "3Vs", [5], [6], [7]: “volume”, relating to the enormous amount of generated data, whether intricately structured or in its raw unstructured form; “velocity”, referring to the rapid and continuous generation of data produced and updated in real-time; “variety”, that indicates the diversity of data types, spanning from the conventional realms of text and numbers to the more intricate landscapes of images and videos. In addition to the “3Vs”, some experts also include other features like “veracity”, [8], referring to data quality and reliability - and “value”, [9], concerning the importance and potential benefit that can be obtained.

Data sets are so massive that traditional data processing software cannot handle them. This has prompted many operators to adopt cloud computing, [10], [11], [12]. It is a consequence of the ease of access to web-based remote computing sites. Instead of installing software for each computer, it only requires installing a simple application on the local device, sending processing data to a single computer in the cloud, [13], [14]. Big data and cloud computing can be considered as a useful and unique tool for many practical applications, [15], [16], [17]. Cloud computing, at its core, is a paradigm that entails delivering computing services-including storage, processing power, and analytics on internet. It is the digital infrastructure that seamlessly supports the processing demands and storage requisites necessitated by the deluge of big data. This dynamic synergy between big data and cloud computing has not only streamlined data management but has also democratized access to advanced computing resources, enabling organizations to leverage the full potential of big data analytics. Therefore, cloud computing not only provides devices for computing and processing big data but also works as a service model, [18], [19]. There are numerous joint applications of big data and cloud computing, for example, in education, [20], [21], [22], agriculture, [23], [24], [25], [26], healthcare, [27], [28], [29], business, [30], [31], [32], [33] and in many other areas, such as transportation, [34], [35].

Data quality, privacy concerns, cybersecurity risks caused by increased data volumes, data interoperability, advanced infrastructure needs, and the requirement for skilled personnel for effective big data analysis are among the new difficulties.

Regarding technologies and tools, the Internet of Things (IoT) comprises devices generating real-time data interconnected for in-depth data processing, [36], [37] and automation across sectors. Additionally, Global Positioning Systems (GPS) provide real-time location data, enhancing efficiency, security, and convenience of mobility and logistics through big data analysis, [38], [39], [40].

Modern advancements in technology have revolutionized data acquisition by providing significant volumes of big data at reduced costs compared to conventional methods. Within this framework, these technological innovations play a crucial role. By leveraging these technological advancements, this study aims to review the main applications of big data in transportation and discuss some noteworthy Italian best practices within the transport sector.

Starting from these considerations the aim of this paper is to review some of the main big data applications (best practices) within the transport sector, highlighting their strengths with a focus on innovative changes and directions observed. Specifically, in the following sections, the multipronged nature of the big data revolution and its transformative impact on defining both present and future trends of the transport sector are presented. Furthermore, two of the main recent Italian big-data applications within the transport sector are discussed, starting from the database of the Italian Ministry of Infrastructure and Transport and the Ministry of Health. The first one investigates the proper national demand estimation by transport mode and territorial area of interest, while the second one correlates the diffusion of the Covid-19 pandemic with the mobility habits in the Country.

The paper is organized as follows: Section 2 describes several big data applications in transportation systems. Section 3 discusses some of the best practices in Italy. Finally, the conclusions and research perspectives are reported in Section 4.

2 Big Data Applications in Transportation Systems

In the vast field of transportation, the integration of big data applications has emerged as a transformative force, offering multifaceted solutions across planning, design, and management domains. These innovations produce substantial advantages, ranging from the optimization of travel times and reduction of road incidents to addressing environmental concerns. The application of big data in the transport sector extends beyond basic optimization, delving into predictive maintenance, smart transportation systems, and informed decision-making, [39]. The utilization of advanced analytics tools and technologies is revolutionizing the way transportation systems operate, enhancing their safety, reliability, and sustainability. Their impact spans various pivotal areas within the transportation sector as, for example:

- real-time traffic monitoring;
- route optimization for commercial vehicles;
- road safety;
- autonomous and connected vehicles;
- travel demand estimation, mobility habits, and transportation planning.

Real-time traffic monitoring harnesses the power of big data analytics by assimilating data from diverse sources, such as traffic cameras, GPS devices, and traffic sensors, enabling the swift identification of traffic congestion and prompt implementation of corrective measures, [40], [41], [42], [43].

Big data facilitates route optimization for freight vehicles, streamlining journeys, curbing costs, and bolstering operational efficiency. Concurrently, it facilitates predictive vehicle maintenance, preempting breakdowns through cloud-based systems monitoring critical components like gearbox oil or clutch discs, [44], [45], [46].

The improvement of road safety is a significant milestone in this path, [47], where big data can be employed to enhance current systems with data-driven warning mechanisms, automatic braking systems, and thorough analysis of accident causation, [48], [49], [50].

The introduction of self-driving and connected vehicles and services (e.g. Google car (2015) and the Robot taxi fleet in California, August 2023), demonstrates a massive utilization of big data, [51], [52], [53]. These vehicles count on a constant inflow of environmental data to drive with safety, achieved through vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) connectivity. The former enables real-time communication among vehicles, interchanging pivotal information, such as traffic status and emergency alerts. By enabling communication between vehicles and roadway infrastructure, the latter facilitates interactive signaling and informative feedback, leading to enhanced traffic flow.

The transportation sector has been able to unlock unprecedented opportunities to forecast and understand mobility demand by integrating big data analytics, [54], [55]. Authorities can make informed decisions and strategically plan thanks to this transformative approach, by examining the challenging dynamics of transportation needs. These improvements in data collection are not restricted to standard methods, as they draw insights from diverse sources (e.g. traffic sensors, GPS systems, public transportation transactions, mobile apps, and social media; [56], [57], [58], [59]. Data sets are combined to provide a complete understanding of travel patterns, vehicle positions, trip durations, and the reasons for individual trips.

A thorough analysis process occurs after data collection, which uncovers intricate mobility patterns and highlights significant trends. The crucial points of travel behavior are revealed during this phase, which includes peak travel hours, routes that are frequently used, and recurring congestion points, [60], [61]. Demand models are created by utilizing this wealth of analyzed data. These models, which are complex in their implementation, consider several aspects that affect travel behavior, such as socioeconomic conditions, geographic locations, mode choices, and available paths. These demand models become very significant in forecasting future mobility needs, [62], [63]. For instance, predictive capabilities derived from big data analytics can help authorities accurately anticipate forthcoming demands, [64], [65], [66]. This foresight allows transportation agencies to plan infrastructures and services more effectively. Examples of application in this sense are, that transport agencies can adapt the timetables of public transport services to meet expected demand at certain times, thereby reducing waiting times but also optimizing resources, thus enhancing the overall efficiency of the transportation ecosystem, [67], [68].

Finally, the integration of big data analytics in transportation planning contributes to sustainable mobility initiatives and revolutionizing the way transportation services are envisioned, planned, and executed.

3 Some Recent Big Data Best Practices Applications in Italy

In the upcoming subsections, the focus will be on presenting successful case studies within the Italian transportation sector, drawing reference from the periodic report "Mobility Trends Observatory" of the Italian Ministry of Infrastructure and Transport, [69], [70], [71] that summarizes the main insights of the big data database on the Italian mobility habits.

Additionally, the main findings of a research study carried out by the Italian University of Campania "Luigi Vanvitelli" and performed by some of the authors of this paper, examining the role of mobility in pandemic diffusion, [72], [73], will also be discussed.

3.1 The Mobility Trends of the Italian Observatory of the Ministry of Infrastructure and Transport

In the spring of 2020, a Mobility Habits Observatory was established by the Ministry of Infrastructures and Transport based on Big Data available from many sources (e.g. transport modes, and distance bands). Its mission also includes disseminating mobility data and statistics within the scientific and scholarly community to bolster the Country's scientific production and deepen knowledge within the mobility and transportation sector. The data and analyses presented are sourced from transport multimodal operators and the Ministry of Infrastructure and Transport's General Directorates, coordinated by the Ministry of Infrastructure and Transport's Technical Mission Structure:

- National Multimodal Operators:
  - ANAS S.p.A.;
  - Motorway Concessionaire Companies
  - Trenitalia S.p.A.;
  - Italo - Nuovo Trasporto Viaggiatori S.p.A.
  - Port Authority;
  - Assaeroporti - Italian Airport Operators Association;
  - FS Research Centre of Ferrovie dello Stato Italiane Group.
- Ministry of Infrastructure and Transport's General Directorates:
  - General Directorate for Roads and Motorways and Oversight and Safety in Road Infrastructures;
  - General Directorate for Oversight of Motorway Concessionaires;
  - General Directorate for Railway Infrastructures and Railway Interoperability;
  - General Directorate for Transport and Railway Infrastructures;
  - General Directorate for Oversight of Port Authorities, Port Infrastructures, Maritime Transport, and Inland Waterways;
  - General Directorate for Airports and Air Transport.

The last published report of the Observatory is the second quarter of 2023, [71]), and Fig. 1 and Fig. 2 report some examples of trend representations, showing, for each mode of transport, the evolution of demand for 2019, 2020, 2021, 2022 and until the second quarter of 2023 (where available). In particular, Fig. 1 reveals that in 2020, during the first wave of the COVID-19 pandemic, a radical decline in mobility demand from February to March was observed. Afterwards, there was a period of recovery, up to August, followed by a reduction until November during the second main wave of COVID-19 diffusion in Italy. Starting from 2021 there was then a slow recovery of demand levels (with different seasonal fluctuations), which led up to 2022 when the mobility levels observed in the pre-pandemic (2019) were (approximately) recovered.

The analyses have been carried out avoiding, where not strictly necessary, any form of data processing, also to minimize possible subjective interpretations of the results.

From these analyses emerges that the large amount of mobility data is useful for estimating mobility habits if they are adequately treated (high professional skills are necessary) to certify the quality of these data. Furthermore, even multi-source and multi-format data can contribute to a better knowledge of the phenomenon, but only if they are adequately archived and processed.

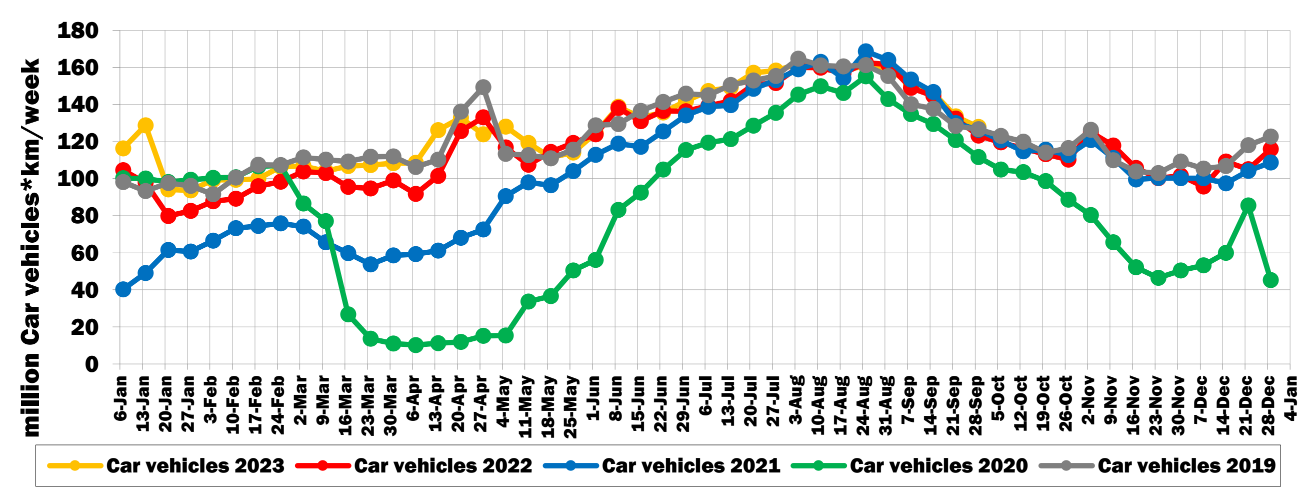

Fig. 1: Car vehicles*km on highways (January 2019-June 2023)
Source: [71]

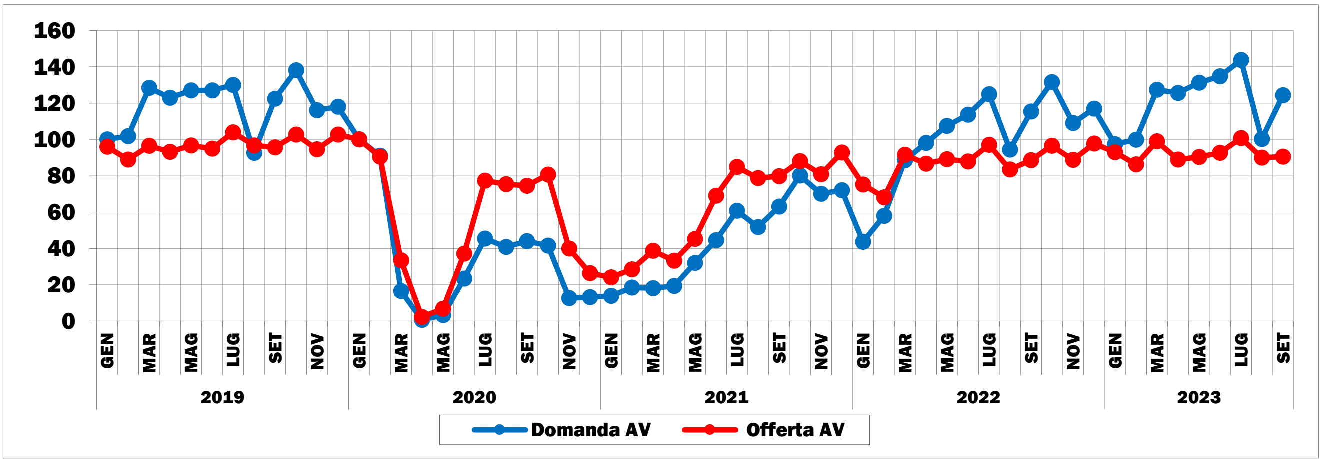

Fig. 2: Monthly demand and supply of high-speed rail passengers (January 2019-June 2023).
Base 100 = value January 2020
Source: [71]

3.2 The Contrast to COVID-19: The Role of Mobility in the Spread of the Pandemic in Italy

The Engineering Department of the University of Campania "Luigi Vanvitelli" conducted a research study to explore the role of mobility in pandemic dissemination and the utilization of big data of the Italian Ministry of Infrastructure and Transport and the Italian Ministry of Health for this aim. This study produced two scientific results published in high-impact international journals, [72], [73].

The COVID-19 pandemic has led to a high number of deaths, economic hardship, and the disruption of daily life. Consequently, the pandemic significantly altered mobility habits and mode choices, leading to a preference among users for private vehicles over public transport, [74], [75], [76].

The findings of the analysis of the first study carried out by [72], probing the impact of urban mobility on COVID-19 spread in Italy revealed that mobility habits constitute one of the variables contributing to COVID-19 infections, along with the daily number of tests conducted and several environmental factors (such as PM particle pollution and temperature), [76]. Notably, regions near the epicenter exhibited a higher infection risk, especially during the initial phase of the outbreak (temporal decay phenomena).

The study, utilizing data from the Official National Monitoring Observatory "Audimob", [77], examines mobility rates in Italy through periodic phone and computer interviews, covering mobility habits pre and post the national lockdown. 2,175 interviews were conducted between January and February (pre-COVID-19) and 1,398 interviews immediately after the lockdown in March 2020. Precisely, the authors, [72], estimate a multiple linear regression model linking daily certified COVID-19 cases to socio-economic, environmental, health care, and mobility habit variables. The study considered daily new positive COVID-19 cases as dependent variables and tested various independent variables at the regional level:

$$Y_{r,i} = \beta_1 \cdot POP_{d,r} + \beta_2 \cdot PM_r + \beta_3 \cdot NTESTS_{r,i} + \beta_4 \cdot TT_{r,i} + \beta_5 \cdot TRIPS_{r,i-x} + \beta_6 \cdot TEM_{r,i-x} + Const \quad (1)$$

where:
$Y_{r,i}$ represents the daily count of new positive COVID-19 cases in the r-th region on the i-th day (source: Italian Ministry of Health, 2020);
$POP_{d,r}$ indicates the population density [10 ∗ inhabitants/km^2] in the provincial capital of the r-th region, [78].
PM_r is the Particulate Matter (PM) pollutant variable [number of days], measuring the days in 2019 when the national PM_{10} daily limit exceeded 50 µg/m^3, [79].
$NTESTS_{r,i}$ is the health care variable estimating the number of COVID-19 tests performed on the i-th day per 1000 population in the r-th region, [70].
$TT_{r,i}$ is the weighted average travel time [hours] from the r-th region to the initial COVID-19 cluster in Codogno (Lombardy) on the i-th day.
$TRIPS_{r,i-x}$ is the average number of people aged 14–80 who made at least one trip (“mobility habits”) “x” days before the i-th day in the r-th region [100,000 ∗ people/day]. This variable investigates the correlation between daily certified coronavirus cases and mobility habits made “x” days earlier.
$TEM_{r,i-x}$ is the average daily temperature “x” days before the i-th day in the r-th region [°C], [80].
$Const$ is a constant variable that was estimated to account for attributes not explicitly covered in the model.

For more details about the model formulation and the hypothesis performed refer to the paper [72].

The daily count of new cases demonstrated a correlation with travels done three weeks earlier, suggesting a potential 21-day period for detecting positivity (Fig. 3). This timeframe implies that the commonly implemented 14-day quarantine based solely on incubation-based epidemiological considerations might underestimate the virus containment approach due to potential delays between infection and detection.

The second study conducted by [73], discussed in this paper aimed to support policymakers and decision-makers in formulating optimal strategies to address the COVID-19 crisis, both from a transportation and security standpoint.

Specifically, the study investigated the correlation between positive COVID-19 cases and the accessibility of transportation within a specific area.

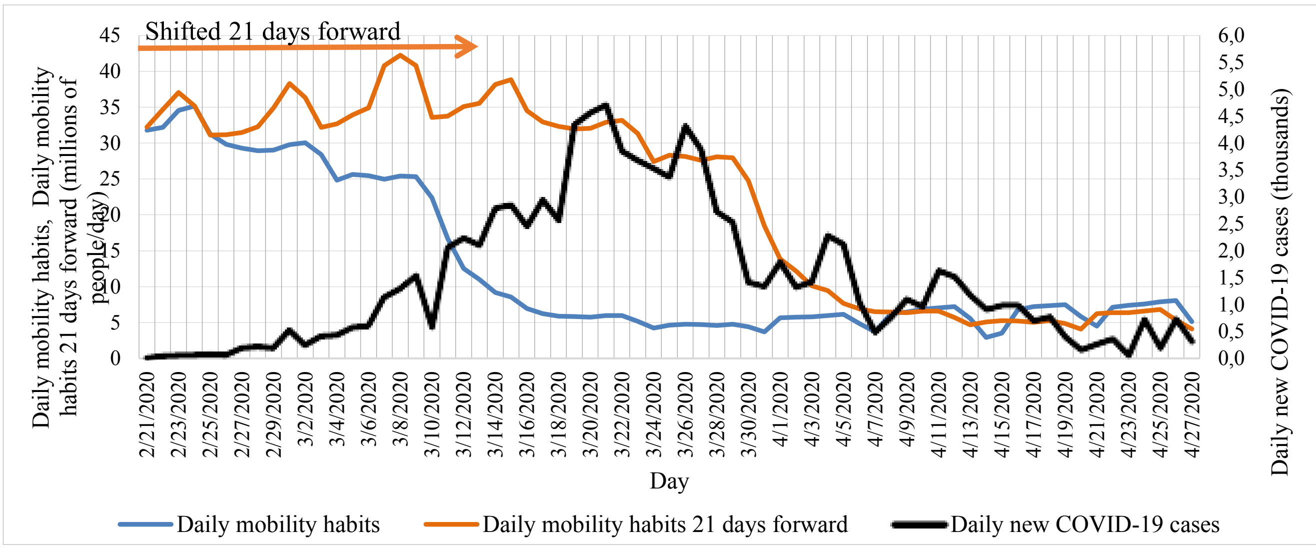

Fig. 3: Delta new COVID-19 cases/day, observed daily 14–80 years old population mobility habits and daily mobility habits shifted 21 days forward (positivity detection time); estimation starting from the big data database of both the Italian Ministry of Health and the Italian Ministry of Infrastructure and Transport
Source: [72]

Estimates were derived through provincial (zonal) aggregation. The data considered encompassed daily reports on new Coronavirus cases (from February 21 to May 5, 2020, sourced from the Italian Ministry of Health), Italian National census data for 2019 (ISTAT), Particulate Matter measurements in 2019 (Italian Regional Environmental Protection Agency - ARPA), daily rail service characteristics from February to May 2020 (Trenitalia, NTV), and an ad-hoc survey in five major rail stations nationwide, conducted from October to November 2019, capturing key mobility habits of rail passengers.

The authors estimated, [73] a multiple linear regression model, linking Italy's total COVID-19 cases to socioeconomic, territorial, and transport accessibility variables:

$$Y_{p,i} = \beta_1 \cdot POP_p + \beta_2 \cdot POP_{d,p} + \beta_3 \cdot SOU_p + \\ + \beta_4 \cdot PM_p + \beta_5 \cdot ACC_p + Const \quad (2)$$

where:
$Y_{p,i}$ is the dependent variable representing the total COVID-19 positive cases in the p-th province from February 21 to April 20, 2020, [70];
POP_p is the population in the p-th province, [78];
$POP_{d,p}$ is the population density in the p-th province, [78].
SOU_p is a dummy variable, equal to 1 for provinces in southern Italy (including Sicily), characterized by warmer weather and extensive coastlines, and 0 otherwise [1; 0];
PM_p measures particulate matter in the p-th province, indicating the number of days in 2019 when the national PM_{10} daily limit exceeded 50 $\mu g/m^3$, [79], [100*number of days];
ACC_p is the proposed rail-based accessibility measure for the p-th province towards all study area zones, calculated using Equation (2) [number/100];
$Const$ is a variable encompassing attributes not otherwise explained in the model [number].
For more details about the model formulation and the hypothesis performed refer to the paper [73].

The findings from estimations revealed that transportation accessibility was the variable that best explained the number of COVID-19 infections (approximately 40% in weight) (Fig. 4). This implies that the higher the accessibility within a specific geographical area, the more easily the virus reaches its population. Furthermore, other contextual factors, such as socioeconomic, territorial, and pollution-related variables were found to be significant. The results suggest that accessibility, typically an indicator of an area's prosperity, becomes a primary conduit for contagion during a pandemic. The quantitative assessments conducted suggest that a potentially more sustainable approach to curbing social interactions could involve tailoring lockdown measures in proportion to the transportation accessibility of the respective areas. This approach proposes that areas with higher accessibility warrant stricter mobility restriction policies, thereby potentially mitigating the spread of the virus more effectively.

Parameters' estimation results of multiple regression models (Model 1, [72] and Model 2, [73]) are shown in Table 1.

The original findings of these two studies have been useful in shaping actions and mobility restriction policies to counter the pandemic. Furthermore, they pave the way for defining potential policies or best practices to enhance the management of mobility restrictions.

4 Conclusion

The emergence of big data has sparked a paradigm shift across various sectors, including transportation, offering a transformative impact on decision-making processes and service delivery. The integration of big data analytics in transportation systems has brought forth a myriad of applications, ranging from real-time traffic monitoring to route optimization, road safety, and even shaping the trajectory of autonomous and connected vehicles. These applications have substantially contributed to optimizing travel times, reducing road incidents, and fostering sustainable mobility initiatives.

In the context of Italy, the Mobility Trends Observatory of the Ministry of Infrastructure and Transport, [71], has played a crucial role in monitoring national mobility demand, contributing valuable insights for infrastructure investments and transportation planning. Additionally, the research conducted by the University of Campania "Luigi Vanvitelli" investigated the role of mobility in the spread of COVID-19. The findings highlighted the significant influence of mobility habits on infection rates and the correlation between transportation accessibility and the rapid spread of the virus. The study findings significantly influenced mobility restriction policies, proposing customized lockdown measures and more effective strategies for managing mobility.

The two case studies discussed in this paper should be considered as possible examples of big-data applications in the transport sector (the topic of this paper). Their applicability and transferability to other case studies (e.g. regions, nations) must be verified case by case and will be the subject of future research.

As we navigate the future, leveraging big data in transportation systems remains pivotal, not only for informed decision-making but also for optimizing service delivery and addressing future challenges in the ever-evolving landscape of mobility and transportation.

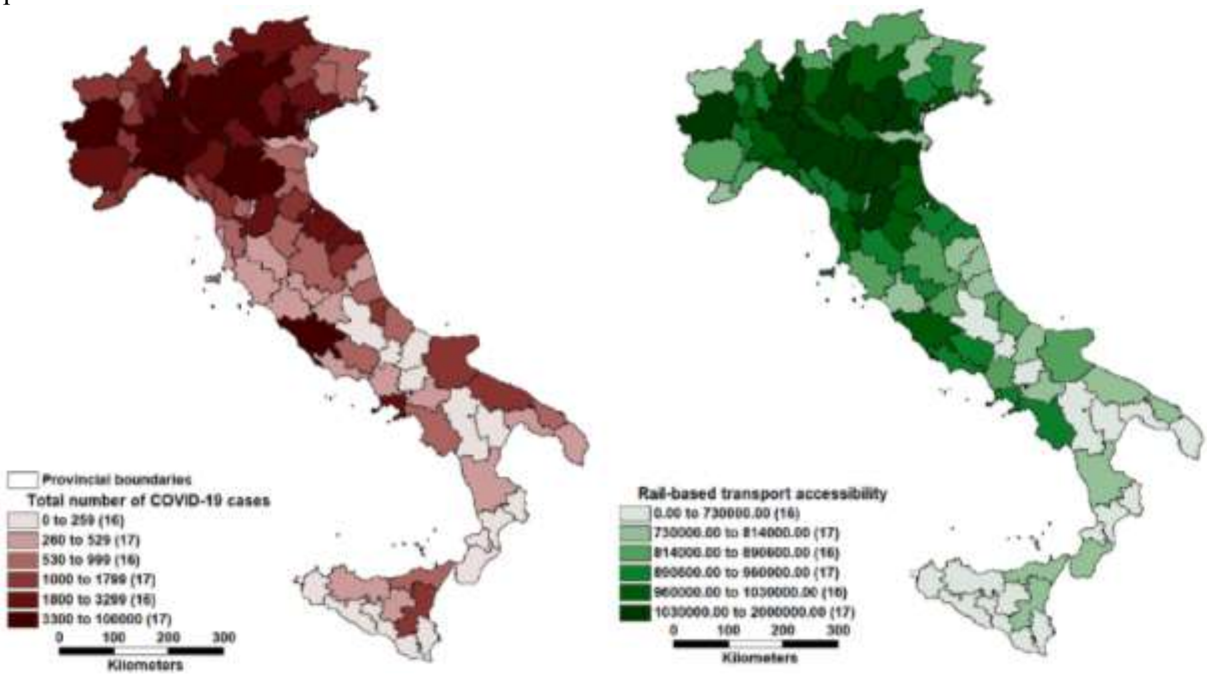

Fig. 4: Total number of COVID-19 cases in Italy (left side - estimation starting from the big data database of the [70]) and rail-based transport accessibility model (2) estimation results (right side - estimation starting from the big data database of the Italian Ministry of Infrastructure and Transport)

Source: [73]

Table 1. Multiple regression models (Model 1 and Model 2): parameters' estimation results

	Model 1 (source: [72])			**Model 2** (source: [73])		
Variable	Std. error	t-value	p-value	Std. error	t-value	p-value
POP_{j*}				0.040	5.057	0.419
$POP_{d,j*}$	0,069	2.299	0.022	0.403	1.850	0.067
PM_{j*}	0.291	2.944	0.003	0.107	2.648	0.009
$NTESTS_{r,j*}$	0.254	7.495	<0.0001			
$TT_{j*,i}$	2.119	-2.336	0.020			
$TRIPS_{j*,i-x}$	0.531	18.460	<0.0001			
$TEM_{j*,i-x}$	1.388	-4.724	<0.0001			
SOU_{j*}				-401.445	1.840	0.069
ACC_{j*}				0.197	2.304	0.023
$Const$ [number]	9.270	1.979	0.048	-1788.444	2.111	0.037
Number of observations			1200			105
R-squared			0.427			0.571
Adj. R-squared			0.424			0.549
F-statistic (6, 1193)			148.359			25.568
P-value (F)			1.30E–140			2.58E–16

**The subscript "j" denotes the region for Model 1, while it refers to the province territorial aggregation when referencing Model 2, where it designates the province.*

Source: [72] and [73]

References:

[1] Saleh, M., & Hamdan, I. T. Analysis on Security Vulnerabilities of Medical Wearable Devices (Fitness Trackers). *WSEAS Transactions on Engineering World*, 2020, 2, pp. 162-166.

[2] Mamedova, N., Belyakova, E., & Urintsov, A. Selection of Organizational Structure of the Company in the Period of Digital Transformation Using the SCARF Model. *Engineering World*, 2023, 5, pp. 29-37, https://doi.org/10.37394/232025.2023.5.4.

[3] Kaisler, S., Armour, F., Espinosa, J. A., and Money, W. Big data: Issues and challenges moving forward. In: *2013 46th Hawaii international conference on system sciences*. IEEE, 2013. pp. 995-1004.

[4] Khalil, M. L., Aziz, N. A., Ariffin, A. A. M., & Ngah, A. H. Big Data Analytics capability and firm performance in the hotel industry: The mediating role of Organizational Agility. *WSEAS Transactions on Business and Economics*, 2023, Vol. 20, pp. 440-453, https://doi.org/10.37394/23207.2023.20.40.

[5] Zikopoulos, P. C., Deroos, D. and Parasuraman, K. *Harness the power of big data: The IBM big data platform*, McGraw-Hill, 2013.

[6] Berman, J. J. *Principles of big data: preparing, sharing, and analyzing complex information*. Newnes, 2013.

[7] Kitchin, R., & McArdle, G. "What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets". *Big Data & Society*, 2016, Vol. 3(1), 2053951716631130.

[8] Rubin, V., & Lukoianova, T. Veracity roadmap: Is big data objective, truthful, and credible? *Advances in Classification Research Online*, 2013, Vol. 24(1), pp. 4.

[9] Gantz, J., & Reinsel, D. Extracting value from chaos. *IDC iview*, 2011, Vol. 1142(2011), pp. 1-12.

[10] Liu, H. Big data drives cloud adoption in enterprise. *IEEE internet computing*, 2013, Vol. 17(4), pp. 68-71.

[11] O'Driscoll, A., Daugelaite, J. and Sleator, R. D. 'Big data', Hadoop and cloud computing in genomics. *Journal of biomedical informatics*, 2013, Vol. 46(5), pp. 774-781.

[12] Behrend, T. S., Wiebe, E. N., London, J. E., & Johnson, E. C. Cloud computing adoption and usage in community colleges. *Behaviour & Information Technology*, 2011, Vol. 30(2), pp. 231-240.

[13] Mathur, P. and Nishchal, N. (2010) 'Cloud computing: New challenge to the entire computer industry. In: *2010 First International Conference On Parallel, Distributed and Grid Computing (PDGC 2010), IEEE*, 2010, pp. 223-228.

[14] Jadeja, Y. and Modi, K. Cloud computing-concepts, architecture and challenges. In: *2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET), IEEE*, Kumaracoil, India, 2012, pp. 877-880.

[15] Yang, C., Huang, Q., Li, Z., Liu, K., & Hu, F. Big Data and cloud computing: innovation opportunities and challenges. *International Journal of Digital Earth*, 2017, Vol. 10(1), pp. 13-53.

[16] Elhoseny, M., Abdelaziz, A., Salama, A. S., Riad, A. M., Muhammad, K., & Sangaiah, A. K. A hybrid model of internet of things and cloud computing to manage big data in health services applications. *Future generation computer systems*, 2018, Vol. 86, pp.1383-1394.

[17] Chong, H. Y., Wong, J. S., & Wang, X. (2014). An explanatory case study on cloud computing applications in the built environment. *Automation in Construction*, 2014, Vol. 44, pp. 152-162.

[18] Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A. and Khan, S. U. The rise of "big data" on cloud computing: Review and open research issues. *Information systems*, 2015, Vol. 47, pp. 98-115.

[19] Talia, D. Clouds for scalable big data analytics. *Computer*, 2013, Vol. 46(5), pp. 98-101.

[20] Singh D. Cloud Computing Environment in Big Data for Education. *Frontiers of Cyberlearning: Emerging Technologies for Teaching and Learning*. 2018, pp. 211-233.

[21] Khan, S., Shakil, K. A. and Alam, M. Educational intelligence: Applying cloud-based big data analytics to the Indian education sector. In: *2016 2nd international conference on contemporary computing and informatics (IC3I), IEEE*, 2016, pp. 29-34.

[22] Baig, M. I., Shuib, L., & Yadegaridehkordi, E. Big data in education: a state of the art, limitations, and future research directions. *International Journal of*

Educational Technology in Higher Education, 2020, Vol.17(1), pp. 1-23.

[23] Rajeswari, S., Suthendran, K. and Rajakumar, K. A smart agricultural model by integrating IoT, mobile and cloud-based big data analytics. In: *2017 international conference on intelligent computing and control (I2C2), IEEE*, 2017, pp. 1-5.

[24] Gill, S. S., Chana, I. and Buyya, R. IoT based agriculture as a cloud and big data service: the beginning of digital India. *Journal of Organizational and End User Computing (JOEUC)*, 2017, Vol. 29(4), pp. 1-23.

[25] Misra, N. N., Dixit, Y., Al-Mallahi, A., Bhullar, M. S., Upadhyay, R., & Martynenko, A. IoT, big data, and artificial intelligence in agriculture and food industry. *IEEE Internet of things Journal*, 2020, Vol. 9(9), pp. 6305-6324.

[26] Liu, W. Smart sensors, sensing mechanisms and platforms of sustainable smart agriculture realized through the big data analysis. *Cluster Computing*, 2023, Vol. 26(5), pp. 2503-2517.

[27] Rajabion, L., Shaltooki, A. A., Taghikhah, M., Ghasemi, A., and Badfar, A. Healthcare big data processing mechanisms: the role of cloud computing. *International Journal of Information Management*, 2019, Vol. 49, pp. 271-289.

[28] Lo'ai, A. T., Mehmood, R., Benkhlifa, E. and Song, H. Mobile cloud computing model and big data analysis for healthcare applications. *IEEE Access*. 2016, Vol. 4, pp. 6171-6180.

[29] Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. Big data in healthcare: management, analysis and prospects. *Journal of big data*, 2019, Vol. 6(1), pp. 1-25.

[30] Balachandran, B. M. and Prasad, S. Challenges and benefits of deploying big data analytics in the cloud for business intelligence. *Procedia Computer Science*, 2017, Vol. 112, pp. 1112-1122.

[31] Wang, Z. and Zhao, H. Empirical study of using big data for business process improvement at a private manufacturing firm in cloud computing. In: *2016 IEEE 3rd international conference on cyber security and cloud computing (CSCloud), IEEE*, 2016, pp. 129-135.

[32] Frizzo-Barker, J., Chow-White, P. A., Mozafari, M., & Ha, D. An empirical study of the rise of big data in business scholarship. *International Journal of Information Management*, 2016, Vol. 36(3), pp. 403-413.

[33] Wiener, M., Saunders, C., & Marabelli, M. Big-data business models: A critical literature review and multiperspective research framework. *Journal of Information Technology*, 2020, Vol. 35(1), pp. 66-91.

[34] Chung, S. H. Applications of smart technologies in logistics and transport: A review. *Transportation Research Part E: Logistics and Transportation Review*, 2021, 153, 102455.

[35] Schumann, H. H., Haitao, H., & Quddus, M. Passively generated big data for micro-mobility: State-of-the-art and future research directions. *Transportation Research Part D: Transport and Environment*, 2023, 121, 103795.

[36] Tang, S., Shelden, D. R., Eastman, C. M., Pishdad-Bozorgi, P., & Gao, X. A review of building information modeling (BIM) and the Internet of things (IoT) devices integration: Present status and future trends. *Automation in Construction*, 2019, Vol. 101, pp. 127-139.

[37] Breivold, H. P., & Sandström, K. Internet of things for industrial automation--challenges and technical solutions. In *2015 IEEE International Conference on Data Science and Data Intensive Systems* IEEE, 2015, pp. 532-539.

[38] Torre-Bastida, A. I., Del Ser, J., Laña, I., Ilardia, M., Bilbao, M. N., & Campos-Cordobés, S. (2018). Big Data for transportation and mobility: recent advances, trends, and challenges. *IET Intelligent Transport Systems*, 2018, Vol. 12(8), pp.742-755.

[39] Roy, A., & Ghosh, D. A modified neural network model for Real-time Driver Drowsiness detection system. *Engineering World*, 2022, 4, pp. 77-84.

[40] Chen, Y. T., Sun, E. W., Chang, M. F., & Lin, Y. B. Pragmatic real-time logistics management with traffic IoT infrastructure: Big data predictive analytics of freight travel time for Logistics 4.0. *International Journal of Production Economics*, 2021, 238, 108157.

[41] Shan, Z., & Zhu, Q. Camera location for real-time traffic state estimation in urban road network using big GPS data.

Neurocomputing, 2015, Vol. 169, pp. 134-143.

[42] Kashinath, S. A., Mostafa, S. A., Mustapha, A., Mahdin, H., Lim, D., Mahmoud, M. A., & Yang, T. J. Review of data fusion methods for real-time and multi-sensor traffic flow analysis. *IEEE Access*, 2021, 9, 51258-51276.

[43] Shi, Q., & Abdel-Aty, M. Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. *Transportation Research Part C: Emerging Technologies*, 2015, Vol. 58, pp. 380-394.

[44] Lakshmanaprabu, S. K., Shankar, K., Rani, S. S., Abdulhay, E., Arunkumar, N., Ramirez, G., & Uthayakumar, J. An effect of big data technology with ant colony optimization based routing in vehicular ad hoc networks: Towards smart cities. *Journal of cleaner production*, 2019, Vol. 217, pp. 584-593.

[45] Zhu, M., Liu, X. Y., Qiu, M., Shen, R., Shu, W., & Wu, M. Y. Traffic big data based path planning strategy in public vehicle systems. In *2016 IEEE/ACM 24th International Symposium on Quality of Service (IWQoS)* (pp. 1-2). IEEE, 2016.

[46] Hang, L., Kang, S. H., Jin, W., & Kim, D. H. Design and implementation of an optimal travel route recommender system on big data for tourists in Jeju. *Processes*, 2018, Vol. 6(8), pp. 133.

[47] Picone, M., Errichiello, A., & Cartenì, A. How Often are ADAS Used? Results of a Car Drivers' Survey. *WSEAS Transactions on Systems*, 2023, pp. 566-577. DOI: 10.37394/23202.2023.22.57.

[48] Hidalgo, D., Lopez, S., Lleras, N., & Adriazola-Steil, C. Original road safety research: Using big data for improving speed enforcement and road safety engineering measures: An application in Bogota, Colombia. *Journal of the Australasian College of Road Safety*, 2018, Vol. 29(2), pp. 12-19.

[49] Sohail, A., Cheema, M. A., Ali, M. E., Toosi, A. N., & Rakha, H. A. Data-driven approaches for road safety: a comprehensive systematic literature review. *Safety Science*, 2023, 158, 105949.

[50] Amin, M. A., Hadouej, S., & Darwish, T. S. Big data role in improving intelligent transportation systems safety: A survey. In *Advances in Internet, Data and Web Technologies: The 7th International Conference on Emerging Internet, Data and Web Technologies (EIDWT-2019, Springer International Publishing*, 2019, pp. 187-199.

[51] Tussyadiah, I. P., Zach, F. J., & Wang, J. Attitudes toward autonomous on demand mobility system: The case of self-driving taxi. In Information and Communication Technologies in Tourism 2017: Proceedings of the International Conference in Rome, Italy, January 24-26, 2017, pp. 755-766. *Springer International Publishing*.

[52] Cascetta, E., Cartenì, A., Di Francesco, L. Do autonomous vehicles drive like humans? A Turing approach and an application to SAE automation Level 2 cars. *Transportation Research Part C: Emerging Technologies*, 2022, 134. DOI: 10.1016/j.trc.2021.103499.

[53] Lian, Y., Zhang, G., Lee, J., & Huang, H.: Review on big data applications in safety research of intelligent transportation systems and connected/automated vehicles. *Accident Analysis & Prevention*, 2021, 146, 105711

[54] Croce, A. I., Musolino, G., Rindone, C., & Vitetta, A. Estimation of travel demand models with limited information: Floating car data for parameters' calibration. *Sustainability*, 2021, Vol. 13(16), 8838.

[55] Cantelmo, G., & Viti, F. A big data demand estimation model for urban congested networks. *Transport and Telecommunication Journal*, 2020, Vol. 21(4), pp. 245-254.

[56] Zhu, L., Yu, F. R., Wang, Y., Ning, B., & Tang, T. Big data analytics in intelligent transportation systems: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 2018, Vol. 20(1), pp. 383-398.

[57] Welch, T. F., & Widita, A. Big data in public transportation: a review of sources and methods. *Transport reviews*, 2019, Vol. 39(6), pp. 795-818.

[58] Pandey, M., Litoriya, R., & Pandey, P. Mobile applications in context of big data: A survey. In *2016 Symposium on Colossal Data Analysis and Networking (CDAN)*, IEEE, 2016, pp. 1-5.

[59] Ghani, N. A., Hamid, S., Hashem, I. A. T., & Ahmed, E. Social media big data analytics: A survey. *Computers in Human behavior*, 2019, Vol. 101, pp. 417-428.

[60] Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C. The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies*, 2015, Vol. 58, pp. 162-177.

[61] Fusco, G., Bracci, A., Caligiuri, T., Colombaroni, C., & Isaenko, N. Experimental analyses and clustering of travel choice behaviors by floating car big data in a large urban area. *IET Intelligent Transport Systems*, 2018, Vol. 12(4), pp. 270-278.

[62] Anda, C., Erath, A., & Fourie, P. J. Transport modeling in the age of big data. *International Journal of Urban Sciences*, 2017, Vol. 21(sup1), pp. 19-42.

[63] Zhao, Y., Zhang, H., An, L., & Liu, Q. Improving the approaches of traffic demand forecasting in the big data era. *Cities*, 2018, Vol. 82, pp.19-26.

[64] Zhang, J., Shen, D., Tu, L., Zhang, F., Xu, C., Wang, Y., ... & Li, Z. A real-time passenger flow estimation and prediction method for urban bus transit systems. *IEEE Transactions on Intelligent Transportation Systems*, 2017, Vol. 18(11), pp. 3168-3178.

[65] Yang, X., Xue, Q., Ding, M., Wu, J., & Gao, Z. Short-term prediction of passenger volume for urban rail systems: A deep learning approach based on smart-card data. *International Journal of Production Economics*, 2021, 231, 107920.

[66] Aqib, M., Mehmood, R., Alzahrani, A., Katib, I., Albeshri, A., & Altowaijri, S. M. Rapid transit systems: smarter urban planning using big data, in-memory computing, deep learning, and GPUs. *Sustainability*, 2019, Vol. 11(10), 2736.

[67] Wang, Y., Zhang, D., Hu, L., Yang, Y., & Lee, L. H. A data-driven and optimal bus scheduling model with time-dependent traffic and demand. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 18(9), pp. 2443-2452.

[68] Liu, W. K., & Yen, C. C. Optimizing bus passenger complaint service through big data analysis: Systematized analysis for improved public sector management. *Sustainability*, 2016, Vol. 8(12), 1319.

[69] Italian Transport Ministry, 2020. Strategic and infrastructures development office: COVID-19 mobility observatory, [Online]. http://www.mit.gov.it/ (Accessed Date: May 7, 2020).

[70] Italian Ministry of Health, 2020. Daily reports on COVID-19 positive cases, [Online]. http://www.salute.gov.it/portale/home.html (Accessed Date: May 7, 2020).

[71] Observatory on passenger and freight mobility trends (III quarter 2023) by the Technical Mission Structure (STM) Cartenì, A., Bazzichelli, T., Carbone, A, [Online]. https://www.mit.gov.it/nfsmitgov/files/media/pubblicazioni/2023-11/REPORT_III%20trimestre%202023.pdf (Accessed Date: November 1, 2023).

[72] Cartenì, A., Di Francesco, L., & Martino, M. How mobility habits influenced the spread of the COVID-19 pandemic: Results from the Italian case study. *Science of the Total Environment*, 2020, 741, 140489.

[73] Cartenì, A., Di Francesco, L., & Martino, M. The role of transport accessibility within the spread of the Coronavirus pandemic in Italy. *Safety science*, 2021a, 133, 104999.

[74] Das, S., Boruah, A., Banerjee, A., Raoniar, R., Nama, S., & Maurya, A. K. Impact of COVID-19: A radical modal shift from public to private transport mode. *Transport Policy*, 2021, 109, 1-11.

[75] Oestreich, L., Rhoden, P. S., da Silva Vieira, J., & Ruiz-Padillo, A. Impacts of the COVID-19 pandemic on the profile and preferences of urban mobility in Brazil: Challenges and opportunities. *Travel Behaviour and Society*, 2023.

[76] Cartenì, A., Di Francesco, L., Henke, I., Marino, T.V., Falanga, A. The role of public transport during the second covid-19 wave in Italy. *Sustainability* (Switzerland), 2021c, Vol. 13(21), art. no. 11905, DOI: 10.3390/su132111905.

[77] Isfort, 2020. 17th Report on the mobility of Italians. Observatory "Audimob" – National Survey on Mobility style and behaviors of Italians (17° Rapporto sulla mobilità degli italiani. Osservatorio "Audimob" – Stili e comportamenti di mobilità degli italiani, 2020), [Online]. https://www.isfort.it/wp-content/uploads/2020/12/RapportoMobilita2020.pdf (Accessed Date: November 1, 2023).

[78] ISTAT - National Institute of Statistics, 2020 (Istituto Nazionale di Statistica, 2020), [Online].

https://www.istat.it/it/archivio/222527 (Accessed Date: November 1, 2023).
[79] ARPA - Regional Agency for Environmental Protection, 2020 (Agenzia Regionale per la Protezione Ambientale, 2020), [Online]. https://www.arpae.it (Accessed Date: November 1, 2023).
[80] ilMeteo, 2020, [Online]. https://www.ilmeteo.it/portale/archivio-meteo (Accessed Date: November 1, 2023).

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors contributed to the present research at all stages from the formulation of the problem to the final findings and solution. In particular, the author

- Antonella Falanga contributed to formal analysis, resources, methodology, data curation, elaboration, validation, writing, review, and editing.
- Armando Cartenì contributed to the conceptualization of the paper, supervision, and review.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflict of interest to declare.