

*AI, Multi-source Data and Advanced Analytics for Socio-Economic and Environmental Systems* brings together a selection of original contributions that explore how artificial intelligence, multi-source data integration, predictive analytics, and advanced statistical methodologies are reshaping the analysis and interpretation of complex socio-economic and environmental phenomena.

The volume stems from the scientific discussions developed during the international conference “*Measuring and Interpreting World Changes with Statistics, Data Science and AI*”, held in Rome from 18 to 20 September 2024. The conference was jointly organised by the Association for Applied Statistics (ASA), the Department of Statistical Sciences of Sapienza University of Rome, and the Italian National Institute of Statistics (ISTAT), with the participation of several academic and institutional partners. The event provided a multidisciplinary forum for examining the contribution of statistics, data science, and artificial intelligence to the understanding of contemporary economic, environmental, and social transformations.

The Special Issue addresses both methodological and applied perspectives, focusing on the increasing integration of administrative archives, business registers, digital platforms, mobility data, social media content, and artificial intelligence systems within analytical and decision-support processes. Particular attention is devoted to multi-source statistical frameworks, predictive models, Large Language Models, mobility and tourism analytics, network analysis, labour market intelligence, and climate-oriented foresight methodologies.

The contributions collected in this volume investigate advanced methodological approaches and empirical applications aimed at improving the interpretation of complex systems and strengthening evidence-based decision-making. By combining statistical rigour with machine learning techniques, spatial intelligence, and AI-assisted analytical systems, the volume highlights both the opportunities and the challenges associated with the growing convergence between official statistics, advanced analytics, and digital infrastructures.

Intended for researchers, practitioners, and policymakers, this Special Issue provides a coherent overview of current developments at the intersection of artificial intelligence, multi-source data integration, and socio-economic analytics, contributing to the ongoing scientific debate on how data-driven systems can support governance, sustainability, and strategic planning in increasingly interconnected environments.

This volume is published within the TESI & TEMI editorial series, jointly promoted by Universitas Mercatorum and the Centro Studi delle Camere di Commercio G. Tagliacarne, and reflects their shared commitment—developed in cooperation with the Association for Applied Statistics—to fostering high-quality research and scientific dialogue in applied statistics and official data analysis.

ISBN 978-88-9326-285-9

Special issue 3 *AI, Multi-source Data and Advanced Analytics for Socio-Economic and Environmental Systems* TEMI

TEMI  
territori  
economie  
mercati  
istituzioni



DIPARTIMENTO  
DI SCIENZE STATISTICHE  
SAPIENZA  
UNIVERSITÀ DI ROMA



Istat  
Istituto Nazionale  
di Statistica

Special issue 3

# AI, Multi-source Data and Advanced Analytics for Socio-Economic and Environmental Systems

Editors

Fabio Crescenzi, Luigi Fabbris, Andrea Mazzitelli, Alessandra Righi,  
Alessandro Rinaldi, Maurizio Vichi

Articles

**Official Business Statistics in Italy: Development, Innovations and Perspectives**

Fabio Massimo Rapiti

**Reducing the Dimensionality of a Set of Short Term Indicators for the Quarterly Estimate of the Value Added in Services Sector in the Quarterly National Accounts: a Comparison of Methodologies**

Marianna Ascione

**Warning Indicators for Business Crisis Prevention: a Statistical Analysis of Lombardy Smes**

Eugenio Brentari, Gloria Ermici, Andrea Zanetti

**Multi-Source Analysis of Labor Demand Dynamics**

Annalisa Lucarelli, Alessandra Righi, Roberta Rizzi

**Advancing Customer Sentiment Quantification and Sales Prediction with Large Language Models**

Domenica Fioredistella Iezzi, Roberto Monte, and Daniele Pasquini

**Mobility and Tourism Analytics: Passive Data, Machine Learning, and LLMS**

Bruno Zamengo, Glauco Mantegari, Denis Cappellari

**Understanding Hotel Cooperation in Venice: a Network Analysis of Key Determinants**

Viviana Amati, Anna Moretti, Rebecca Micol Finzi

**Twin Transition, Smart Specialisations and New Geographies of Innovation**

Sergio Salamone, Francesco Giovanni Truglia

**The Role of Artificial Intelligence in Supporting Human Decisions: Digital Twin for Streamlining Market Research**

Lino Coscione, Cinzia Paterlini, Salvatore Fargnoli

**Counteracting Future Climate Challenges Using Artificial Intelligence Models: an Irish Case Study**

Yuri Calleo, Francesco Pilla, Simone Di Zio



CENTRO STUDI DELLE  
CAMERE DI COMMERCIO  
GUGLIELMO TAGLIACARNE

Universitas  
Mercatorum  
Università telematica delle  
Camere di Commercio Italiane

**TEMI** Territori  
Economie  
Mercati  
Istituzioni



DIPARTIMENTO  
DI SCIENZE STATISTICHE

**SAPIENZA**  
UNIVERSITÀ DI ROMA



Special issue 3

# **AI, Multi-source Data and Advanced Analytics for Socio-Economic and Environmental Systems**



CENTRO STUDI DELLE  
CAMERE DI COMMERCIO  
GUGLIELMO TAGLIACARNE



Università telematica delle  
Camere di Commercio Italiane

EDITORIAL BOARD - SPECIAL ISSUE:

Fabio Crescenzi, Luigi Fabbris, Andrea Mazzitelli, Alessandra Righi, Alessandro Rinaldi, Maurizio Vichi

SCIENTIFIC DIRECTION:

Giovanni Cannata (Rector, Universitas Mercatorum) and  
Gaetano Fausto Esposito (Director General, Centro Studi delle Camere di Commercio G. Tagliacarne)

EDITORIAL OFFICE: Annamaria Jannuzzi

COVER DESIGN: Giapeto Editore srl con socio unico - Napoli

EDITORS-IN-CHIEF:

Giovanni Cannata, Gaetano Fausto Esposito

THE JOINT DIGITAL EDITORIAL SERIES PROMOTED BY UNIVERSITAS MERCATORUM AND THE CENTRO STUDI DELLE CAMERE DI COMMERCIO G. TAGLIACARNE INCLUDE:

TESI (Territory, Economy, Society, Institutions). Instant Paper: blog-based publications subject to a preliminary assessment of scientific coherence;

TESI (Territory, Economy, Society, Institutions). Paper: aperiodic publications without ISBN, reviewed through a single-blind peer review process;

TESI (Territory, Economy, Society, Institutions). Discussion Paper: aperiodic publications with ISBN assigned by Universitas Mercatorum, subject to double-blind peer review;

TEMI (Territory, Economy, Markets, Institutions): a series collecting theoretical and analytical contributions selected through thematic calls for papers addressing topics relevant to the scientific communities of Universitas Mercatorum and the Centro Studi delle Camere di Commercio G. Tagliacarne.

*This work, including all of its parts, is protected under applicable copyright law. Any reproduction, distribution, communication, adaptation, translation, or processing for commercial purposes, by any means or formats, including digital platforms, is prohibited without prior authorization. Non-commercial reproduction is permitted provided that the source is properly cited. By downloading this publication, users accept the conditions stated herein.*

DISTRIBUTION PLATFORMS:

[https://www.tagliacarne.it/tesi\\_temi-30](https://www.tagliacarne.it/tesi_temi-30)

<https://www.unimerceatorum.it/ricerca/tesi-e-temi>

APERIODIC PUBLICATION. COPYRIGHT © 2022 PROPRIETORS AND PUBLISHERS:

*Centro Studi delle Camere di Commercio G. Tagliacarne, Universitas Mercatorum*

*Piazza Mattei 10, 00186 Rome*

*Centro Studi delle Camere di Commercio G. Tagliacarne*

*Piazza Sallustio 9, 00187 Rome*

Editor: Giapeto Editore srl con socio unico - Napoli

First edition: June 2026

ISBN: 978-88-9326-285-9

## INDEX

<b>EDITORIAL</b> .....	5
<i>Fabio Crescenzi, Luigi Fabbris, Andrea Mazzitelli, Alessandra Righi, Alessandro Rinaldi, Maurizio Vichi</i>	
<b>LE STATISTICHE UFFICIALI SULLE IMPRESE IN ITALIA: SVILUPPI, INNOVAZIONE E PROSPETTIVE</b> .....	9
<i>OFFICIAL BUSINESS STATISTICS IN ITALY: DEVELOPMENT, INNOVATIONS AND PERSPECTIVES</i>	
<i>Fabio M. Rapiti</i>	
<b>LA RIDUZIONE DELLA DIMENSIONALITÀ DI UN SET DI INDICATORI CONGIUNTURALI PER LA STIMA TRIMESTRALE DEL VALORE AGGIUNTO DEI SERVIZI NEI CONTI ECONOMICI TRIMESTRALI: METODOLOGIE A CONFRONTO</b> .....	25
<i>REDUCING THE DIMENSIONALITY OF A SET OF SHORT TERM INDICATORS FOR THE QUARTERLY ESTIMATE OF THE VALUE ADDED IN SERVICES SECTOR IN THE QUARTERLY NATIONAL ACCOUNTS: A COMPARISON OF METHODOLOGIES</i>	
<i>Marianna Ascione</i>	
<b>INDICATORI DI ALLERTA PER LA PREVENZIONE DELLE CRISI D'IMPRESA: UN'ANALISI STATISTICA SULLE PMI LOMBARDE</b> .....	41
<i>WARNING INDICATORS FOR BUSINESS CRISIS PREVENTION: A STATISTICAL ANALYSIS OF LOMBARDY SMEs</i>	
<i>Eugenio Brentari, Gloria Ermici, Andrea Zanetti</i>	
<b>MULTI-SOURCE ANALYSIS OF LABOR DEMAND DYNAMICS</b> .....	51
<i>ANALISI MULTI-FONTE DELLA DINAMICA DELLA DOMANDA DI LAVORO</i>	
<i>Annalisa Lucarelli, Alessandra Righi, Roberta Rizzi</i>	

<b>QUANTIFICAZIONE AVANZATA DEL SENTIMENT DEI CLIENTI E PREVISIONE DELLE VENDITE MEDIANTE LARGE LANGUAGE MODELS .....</b>	<b>67</b>
<i>ADVANCING CUSTOMER SENTIMENT QUANTIFICATION AND SALES PREDICTION WITH LARGE LANGUAGE MODELS</i>	
<i>Domenica Fioredistella Iezzi, Roberto Monte, and Daniele Pasquini</i>	
<b>ANALISI DELLA MOBILITÀ E DEL TURISMO: DATI PASSIVI, MACHINE LEARNING E LLM .....</b>	<b>85</b>
<i>MOBILITY AND TOURISM ANALYTICS: PASSIVE DATA, MACHINE LEARNING, AND LLMS</i>	
<i>Bruno Zamengo, Glauco Mantegari, Denis Cappellari</i>	
<b>UNDERSTANDING HOTEL COOPERATION IN VENICE: A NETWORK ANALYSIS OF KEY DETERMINANTS .....</b>	<b>103</b>
<i>COMPRENDERE LA COOPERAZIONE TRA HOTEL A VENEZIA: UN'ANALISI DI RETE DEI FATTORI CHIAVE</i>	
<i>Viviana Amati, Anna Moretti, Rebecca Micol Finzi</i>	
<b>TWIN TRANSITION, SPECIALIZZAZIONI INTELLIGENTI E NUOVE GEOGRAFIE DELL'INNOVAZIONE .....</b>	<b>119</b>
<i>TWIN TRANSITION, SMART SPECIALISATIONS AND NEW GEOGRAPHIES OF INNOVATION</i>	
<i>Sergio Salamone, Francesco Giovanni Truglia</i>	
<b>THE ROLE OF ARTIFICIAL INTELLIGENCE IN SUPPORTING HUMAN DECISIONS: DIGITAL TWIN FOR STREAMLINING MARKET RESEARCH .....</b>	<b>143</b>
<i>IL RUOLO DELL'INTELLIGENZA ARTIFICIALE NEL SUPPORTARE LE DECISIONI UMANE: LE DIGITAL TWIN PER SEMPLIFICARE LE RICERCHE DI MERCATO</i>	
<i>Lino Coscione, Cinzia Paterlini, Salvatore Fargnoli</i>	
<b>COUNTERACTING FUTURE CLIMATE CHALLENGES USING ARTIFICIAL INTELLIGENCE MODELS: AN IRISH CASE STUDY .....</b>	<b>155</b>
<i>CONTRASTARE LE SFIDE CLIMATICHE FUTURE UTILIZZANDO MODELLI DI INTELLIGENZA ARTIFICIALE: UN CASO DI STUDIO IRLANDESE</i>	
<i>Yuri Calleo, Francesco Pilla, Simone Di Zio</i>	

## EDITORIAL

*Fabio Crescenzi<sup>1</sup>, Luigi Fabbris<sup>2</sup>, Andrea Mazzitelli<sup>3</sup>, Alessandra Righi<sup>4</sup>,  
Alessandro Rinaldi<sup>5</sup>, Maurizio Vichi<sup>6</sup>*

### 1. Introduction

Official statistics and applied quantitative analysis are currently experiencing a profound transformation driven by the simultaneous expansion of administrative archives, digital infrastructures, georeferenced information systems and artificial intelligence technologies. The increasing availability of heterogeneous and high-frequency data sources has profoundly altered the informational environment within which statistical institutes, public administrations and private organisations operate.

At the same time, global socio-economic systems are becoming increasingly interconnected and dynamic. Climate change, digitalisation, geopolitical instability, labour market transformations, sustainable mobility and changing consumer behaviour require analytical systems capable not only of describing phenomena *ex post*, but also of anticipating trends, simulating scenarios and supporting complex decision-making processes.

In this evolving context, the traditional boundaries between official statistics, data science, predictive analytics and artificial intelligence are progressively becoming more permeable. Statistical systems are moving towards integrated multi-source architectures, while machine learning and Large Language Models are expanding the capacity to extract information from structured and unstructured data. Spatial analytics and network-based approaches are simultaneously strengthening the territorial dimension of socio-economic analysis.

The contributions collected in this special issue reflect these transformations from complementary perspectives. The proposed reading pathway moves from the evolution of official statistics and multi-source systems to predictive analytics and AI-assisted decision frameworks, and finally towards territorial, tourism and climate-oriented applications. The common thread linking the papers is the attempt to combine methodological innovation with interpretability, institutional robustness and policy relevance.

---

<sup>1</sup> Former Istat - Istituto Nazionale di Statistica, Rome, Italy - e-mail: fabio7826@gmail.com

<sup>2</sup> University of Padua, Padua, Italy - e-mail: fabbris@stat.unipd.it

<sup>3</sup> Universitas Mercatorum, Rome, Italy - e-mail: a.mazzitelli@unimercatorum.it

<sup>4</sup> Istat - Istituto Nazionale di Statistica, Rome, Italy - e-mail: righi@istat.it

<sup>5</sup> Centro Studi Tagliacarne, Rome, Italy - e-mail: alessandro.rinaldi@tagliacarne.it

<sup>6</sup> Sapienza University of Rome, Rome, Italy - email: maurizio.vichi@uniroma1.it

## 2. Innovation in Official Statistics

Rapiti in “Official Business Statistics in Italy: Developments, Innovations and Perspectives” analyses the ongoing transformation of Italian business statistics within the framework of the European Business Statistics Regulation. The paper highlights the strategic role of integrated business registers, administrative data and e-invoicing systems in supporting the transition towards a register-based statistical system. The contribution also discusses the institutional and methodological challenges associated with reducing response burden, improving timeliness and expanding the use of alternative data sources.

Ascione in “La riduzione della dimensionalità di un set di indicatori congiunturali per la stima trimestrale del valore aggiunto dei servizi nei Conti economici trimestrali: metodologie a confronto” examines methodological innovations introduced in the quarterly estimation of value added for the services sector within the Italian Quarterly National Accounts. By comparing variable selection and dimensionality reduction techniques, including shrinkage methods and component-based approaches, the paper addresses the problem of identifying optimal predictors among hundreds of short-term indicators. The contribution illustrates how advanced statistical methodologies can improve temporal disaggregation procedures in contexts characterised by fragmented and heterogeneous information.

Lucarelli, Righi and Rizzi in “Multi-source Analysis of Labor Demand Dynamics” explore the integration of administrative, traditional and nontraditional statistical sources to improve the interpretation of labour demand trends in Italy. By comparing Compulsory Communications data, Job Vacancy Surveys and Excelsior survey estimates, the authors analyse the convergence and divergence of labour market indicators across sectors and time frequencies. The study highlights the complementarity of different data sources and the relevance of integrated analytical frameworks for evidence-based labour market policies.

## 3. Predictive Analytics, LLMs and AI-assisted Decision Systems

Brentari, Ermici and Zanetti in “Indicatori di allerta per la prevenzione delle crisi d’impresa: un’analisi statistica sulle PMI lombarde” analyse predictive indicators for business crises in the context of the Italian insolvency reform framework. By comparing the indicators proposed by the National Council of Chartered Accountants with the Altman Z-score model, the study investigates the capacity of statistical indicators to anticipate financial distress among SMEs. The contribution reinforces the role of predictive analytics in supporting organisational resilience and economic governance.

Iezzi, Monte and Pasquini in “Advancing Customer Sentiment Quantification and Sales Prediction with Large Language Models” explore the use of Large Language Mo-

dels for sentiment analysis and sales forecasting. Comparing CardiffNLP, ChatGPT-4 and LLaMA on a corpus of over 160,000 tweets related to the Toyota Camry, the study evaluates the consistency and predictive capacity of different sentiment extraction paradigms. By integrating sentiment proxies within a Multivariate AutoRegressive State-Space model, the paper demonstrates how LLM-generated signals can improve forecasting accuracy compared to traditional approaches.

Coscione, Paterlini and Fagnoli in “The role of artificial intelligence in supporting human decisions: Digital Twin for streamlining market research” investigate the use of Digital Twin technologies in market research and decision support. Combining survey data, social listening and natural language processing techniques, the proposed framework generates realistic customer twins capable of simulating qualitative and quantitative scenarios. The contribution highlights both the opportunities and the ethical implications of AI-assisted decision systems, emphasising the continuing importance of human supervision.

#### **4. Mobility, Tourism and Networked Territorial Systems**

Zamengo, Mantegari and Cappellari in “Enhancing mobility and tourism through data analytics and generative AI” analyse the use of mobile network data, floating car data and generative AI systems for mobility and tourism analytics. The paper illustrates how integrated big data infrastructures can support urban planning, tourism management and sustainable mobility policies through the analysis of visitor flows, transport behaviours and territorial accessibility. The contribution also highlights the growing role of Large Language Models in democratising access to complex analytical systems.

Amati, Moretti and Finzi in “Understanding hotel cooperation in Venice: A network analysis of key determinants” investigate formal and informal collaboration networks among hotels in Venice through weighted Exponential Random Graph Models. The study analyses how geographical proximity, resource endowment and managerial characteristics influence cooperation patterns within a tourism destination characterised by overtourism pressures. The paper demonstrates how network analysis can support sustainable territorial governance and destination management.

Salamone and Truglia in “Twin transition, specializzazioni intelligenti e nuove geografie dell’innovazione” explore the relationship between smart specialisation strategies, digital transition and environmental sustainability across Italian territories. Using spatial correlation techniques and composite indicators, the study identifies territorial clusters and spillover effects associated with smart specialisation domains. The contribution highlights how digitalisation, innovation and sustainability increasingly interact within regional development processes.

## **5. Climate Resilience, Spatial Foresight and Future-oriented Analytics**

Coscione, Paterlini and Fagnoli in “The role of artificial intelligence in supporting human decisions: Digital Twin for streamlining market research” also contribute to a broader reflection on the role of AI-generated simulation environments in supporting anticipatory governance and strategic foresight. Their Digital Twin approach demonstrates how generative systems can create virtual behavioural environments capable of testing alternative decision scenarios before real-world implementation.

Calleo, Pilla and Di Zio in “Counteracting future climate challenges using artificial intelligence models: an Irish case study” propose an innovative integration between Real-Time Spatial Delphi methodologies and generative artificial intelligence models for climate adaptation planning. Applied to an Irish case study involving experts from Cork, Galway and Limerick, the framework combines collaborative spatial decision-making with text-to-image systems capable of visualising future climate scenarios. The contribution illustrates how AI-enhanced foresight methodologies can strengthen the accessibility and communicability of complex environmental analyses for policy makers.

## **6. Concluding reflections: towards integrated cognitive infrastructures**

The ten contributions portray a statistical and analytical ecosystem evolving towards increasing integration between official statistics, artificial intelligence, spatial intelligence and advanced predictive systems. The common thread linking these works is not merely technological innovation, but the attempt to build analytical frameworks capable of combining methodological robustness, interpretability and policy relevance.

The papers collected in this issue show that multi-source systems are progressively transforming statistical production processes, while AI and Large Language Models are expanding the capacity to analyse unstructured information, simulate behavioural dynamics and support anticipatory governance. At the same time, spatial and network-based approaches are reinforcing the territorial dimension of socio-economic and environmental analysis.

In this perspective, official statistics are increasingly evolving from systems primarily devoted to measurement towards integrated cognitive infrastructures capable of supporting strategic decision-making in complex environments. The convergence between statistical science, artificial intelligence and advanced analytics therefore represents not a temporary technological transition, but a structural redefinition of how contemporary societies observe, interpret and govern economic, social and environmental systems.

## LE STATISTICHE UFFICIALI SULLE IMPRESE IN ITALIA: SVILUPPI, INNOVAZIONI E PROSPETTIVE

### OFFICIAL BUSINESS STATISTICS IN ITALY: DEVELOPMENT, INNOVATIONS AND PERSPECTIVES

*Fabio M. Rapiti<sup>1</sup>*

#### **Sommario**

In Italia, le statistiche sulle imprese sono prodotte e diffuse principalmente dall'Istat, l'Istituto Nazionale di Statistica, che attraverso la combinazione di registri statistici integrati delle imprese e rilevazioni multiscopo qualitative fornisce agli utilizzatori un quadro informativo ampio e coerente. I recenti sviluppi e le innovazioni sono stati in parte guidati dai requisiti del Regolamento europeo sulle statistiche sulle imprese (EBS) del 2019, con progressi particolarmente rilevanti nel settore dei servizi. Gli Istituti Nazionali di Statistica affrontano oggi sfide significative, quali la riduzione dei tassi di risposta e la limitatezza delle risorse, a fronte di una crescente domanda di dati in un contesto globale sempre più incerto. La risposta principale consiste nel ricorso a fonti informative e metodologie innovative, che consentono di ridurre le indagini dirette e il carico statistico sulle imprese. Sebbene in molti Paesi lo sfruttamento delle fonti amministrative e dei big data proceda lentamente, l'Istat ha avviato da tempo la transizione verso un sistema basato su registri, facendo ampio uso dei dati amministrativi. Un'evoluzione recente e particolarmente rilevante è rappresentata dai dati di fatturazione elettronica dell'Agenzia delle Entrate, una nuova fonte ad alto valore aggiunto con un elevato potenziale per le statistiche economiche e i conti nazionali, soprattutto in termini di miglioramento della qualità e riduzione dell'onere statistico. Il piano pluriennale di sostituzione e integrazione delle tre rilevazioni congiunturali sul fatturato dell'industria, dei servizi e del commercio al dettaglio ne è un esempio emblematico e si configura come la più ampia riduzione del carico statistico sulle imprese mai realizzata a livello nazionale. Per superare le criticità che ostacolano il pieno sfruttamento delle fonti alternative è necessaria una chiara visione strategica. L'esperienza maturata dall'Istat indica alcune linee di azione fondamentali: l'ampliamento dell'uso di fonti alternative ad alto valore aggiunto, in particolare per i dati quantitativi; la conseguente riduzione delle indagini dirette, preservando quelle necessarie a raccogliere informazioni qualitative (come la rilevazione multiscopo) non altrimenti disponibili; la rimozione delle barriere istituzionali

---

<sup>1</sup> Istat - Istituto Nazionale di Statistica, Italia - e-mail: [fabio.rapiti@istat.it](mailto:fabio.rapiti@istat.it)

e operative all'accesso tempestivo ai dati amministrativi; il rafforzamento della condivisione dei microdati; e il consolidamento della cooperazione inter-istituzionale e internazionale.

### **Abstract**

*In Italy, business statistics are produced and disseminated primarily by Istat, the National Statistical Institute, which, through the combination of integrated statistical business registers and qualitative multi-purpose surveys, provides users with a comprehensive and coherent information framework. Recent developments and innovations have been partly driven by the requirements of the 2019 European Business Statistics (EBS) Regulation, with particularly significant progress in the services sector.*

*National Statistical Institutes currently face major challenges, including declining response rates and constrained resources, against a backdrop of growing demand for data in an increasingly uncertain global environment. The main response has been a shift towards innovative data sources and methodologies, enabling a reduction in direct surveys and in the statistical burden on firms. While in many countries the full exploitation of administrative sources and big data is progressing slowly, Istat has long embarked on the transition towards a register-based statistical system, making extensive use of administrative data. A recent significant development is the availability of e-invoicing data from the Italian Revenue Agency (Agenzia delle Entrate), a new high value-added administrative source with considerable potential for economic statistics and national accounts, especially in terms of quality improvements and reductions in response burden. The multi-year plan to replace and integrate the three short-term surveys on turnover in industry, services and retail trade constitutes a clear example of this approach and represents the most substantial reduction in the statistical burden on enterprises ever achieved at the national level. To overcome the challenges that hinder the full exploitation of alternative data sources, a clear strategic vision is required. The experience gained by Istat points to several key lines of action: expanding the use of high value-added alternative sources, particularly for quantitative data; the corresponding reduction of direct surveys, while preserving those necessary to collect qualitative information (such as multi-purpose surveys) that cannot be obtained otherwise; the removal of institutional and operational barriers to timely access to administrative data; enhancing microdata sharing, and strengthening inter-institutional and international cooperation.*

**Keywords:** *Business statistics, administrative data, big data, official statistics.*

## 1. Introduction<sup>2</sup>

Business and trade statistics provide information on businesses whose sector contributes 63.2% to GDP (Gross Domestic Product) in Italy (2023). They explain the structure and evolution of economies and provide many of the input data required by the System of National Accounts and the calculation of GDP. They also contribute to answer some of the most important policy questions of our age, for example, how globalization and digitalization are impacting economies, societies, and environments.

Every advanced country maintains official business statistics that enable institutions, economic operators, and citizens to get evidence-based decision-making in both economic and social domains. These statistics are primarily managed by National Statistical Institutes (NSIs). In Italy, as in the rest of the European Union, business statistics are governed mainly by EU regulations and recently have undergone significant development, particularly following the adoption of the 2019 Regulation. Istat completed the full implementation of this Regulation in 2024, addressing several information gaps and notably expanding and improving business statistics in the service sector.

The recent global economy development linked to health and recurring geo-political crises has increasing uncertainty and created new data needs. But despite some difficulties there are opportunities for NSIs to respond appropriately and meet these needs.

At national level, Istat stands out for its strategy of combining integrated business registers and multipurpose surveys, which form the backbone of the Permanent Census (Censimento Permanente). This approach offers users a variety of additional and complementary information – both quantitative and qualitative – compared to what is produced under the EU regulatory framework. Multipurpose surveys, for instance, allow for the collection of different data to support modern analytical needs.

However, recent challenges, including the decline in response rates, the need to reduce the statistical burden on respondents, the cuts in resources allocated to statistics, are forcing NSIs to rethink the forms of data collection and undertake important changes. Among these, the implementation of innovative approaches based on a greater and more effective use of new data sources, especially administrative and big data, with the prospect of reducing traditional direct surveys (Snijkers *et al.*, 2023). In Italy, this transition began many years ago, and Istat has since developed a register-based business statistics system making extensive use of administrative sources, for which a specific quality framework has been established (Istat, 2016). Nevertheless, significant challenges re-

---

<sup>2</sup> This article is an expanded and refined version of the presentation delivered by the author, as Chair, serving as the introductory address to the opening session “Innovations and developments in business statistics: new sources, processes and topics” at the international ASA conference “Measuring and Interpreting World Changes with Statistics, Data Science and AI” held in Rome on 18 September 2024.

main, particularly in further reducing the statistical burden on businesses. A new source that will allow a huge development of statistics and a strong reduction of the burden on firms is e-invoicing. But the full use of alternative sources remains a topic that deserves further study.

This article is divided into 6 parts. In section 2, the recent developments and innovations in Istat business statistics are briefly illustrated, especially in reference to the EU regulatory framework. Section 3 examines in depth the key role that the Multipurpose Survey has in monitoring emerging phenomena and allowing multidimensional analysis. Section 4 describes the current context for business statistics with opportunities and challenges that official statistics providers must address in order to remain relevant in a fast-changing world and to satisfy the new user's needs. Some considerations are presented on the slow progress in the use of administrative data and big data. In the 5th section, the use of the Electronic Invoicing data of the Revenue Agency (Agenzia delle Entrate) is discussed, a high value/high-impact source with enormous potential for all economic statistics. In particular, the work in progress in short-term turnover statistics is presented. Section 6 attempts draws some conclusions providing key actions to move forward the near future.

## **2. Recent developments related to the European Business Statistics Regulation**

Before delving into the recent development in Italian business statistics, it is worth briefly describing how they are produced at a national level. For several years now, a lot of business statistics, including those that are transmitted to Eurostat to comply with the EU Regulations, have been produced using a multi-source approach, combining traditional surveys with administrative and alternative data sources to enhance coverage, timeliness, and data quality.

The Istat business statistics production system is based on a set of core national infrastructures and a mix of instruments and data sources. The main elements are: 1) the Business Portal, a tool where the direct collection of data is centralized in a single place for the relationship with the respondents, where all the surveys (CAWI) and questionnaires are found; 2) the system of SBRs (Statistic Business Registers) – part of a larger Istat Integrated Register Ecosystem – divided into several “basic” and “extended” registers built mainly on the integration of administrative sources; 3) the large triennial Multi-Purpose business Survey (MPS) which, together with the SBRs, constitutes the so-called “Permanent Business Census”.

Business statistics have been heavily regulated at EU level since the second half of the 1990s. The EU Framework Regulation for European Business Statistics-EBS (n. 2152/2019) approved in December 2019 (European Parliament and the Council, 2019) together with the subsequent implementing regulations approved between 2020-2022,

consolidated and harmonised many previous legislative acts. It was the end point of a long process of modernisation of official EU statistics to keep pace with the growing demand for information and the rapid changes in the economy and society. This framework required all member countries to raise significantly the bar by broadening the coverage and detail available for many indicators and obliged them to adapt to high standards in quantitative and qualitative terms. Official statistics, in fact, must follow standards, transparent rules and rigorous methodologies codified by the principles of the European Statistics Code (Eurostat, 2017) and which must be guaranteed by independent, reliable, public statistical institutions, generally the NSIs.

Many member countries, including Italy, requested some years of derogations in order to reach the challenging targets and comply with the Regulation. Istat achieved full compliance in 2024 and some of the most significant and recent innovations implemented are briefly presented below.

For all the short-term surveys on the levels of economic activity in volume and value, the transition to the new base (2021) has been completed. An increase in output has been achieved, especially in the services sector. Since the release referring to the January 2024 data, the turnover indices (Istat, 2024a):

- are available on a monthly basis (previously they were quarterly) and released at approximately t+59 days;
- the data in volume as well as the seasonally adjusted ones are released for all sectors (previously there were no volumes and only some series were seasonally adjusted);
- sector coverage has increased and more time-series are released.

The transition to the new reference base has been also implemented for the three producer price indices (industry, construction and services) and import prices. In July 2024, the statistical information available for the analysis of the dynamics of producer prices of services increased:

- the sectoral coverage of the producer price indices of Business-to-Business (BtoB) services was increased and
- the dissemination of the producer price indices of Business-to-All (BtoAll) services, i.e. relating to the total, business and consumer market, has started (Istat, 2024b).

Furthermore, the production of the new annual statistics on international trade in services by enterprise characteristics (STEC-Services Trade by Enterprise Characteristics) has started (Istat-Ice, 2025). These statistics provide information on enterprises involved in international trade in services and show how enterprises in different sectors pro-

vide services and how this is linked to their main economic activity. For the production of the STEC statistics, a close collaboration with the Bank of Italy has been activated to reuse the data from the TTN-Transazioni Trimestrali Non finanziarie (Non-Financial Quarterly Transactions) Survey – whose sample size has been specifically increased – together with other sources. This ensures greater coherence with the domain of Balance of Payments statistics and has made it possible to avoid the launch of a new business survey, generating further statistical burden.

Since 2023, the monthly system for the exchange of microdata relating to intra-EU exports (MDE - Micro Data Exchange) between the 27 NSIs of the EU member states has been fully operational for Italy. The system enables the reuse of export data collected by Member States by the importing country, in line with the “once only” principle. The exchange of microdata is an innovative approach to reduce the statistical burden of the Intrastat system and provide member countries with an additional and detailed source for compiling statistics on intra-EU imports. Currently, Istat uses the microdata provided by other EU countries as an additional source for compiling statistics on imports to increase accuracy, reduce asymmetries and, together with other fiscal sources, improve estimation processes (non-response and reporting to the universe) (Istat, 2025a). But the conditions are now in place for a substantial reduction, within one years, in the obligation to complete Intrastat forms for broad segments of exporters also thanks to the use of new sources such as e-invoice data.

### **3. The Multi-purpose Business Survey and some examples of innovative contents**

At national level, the business Permanent Census has a central role. It is based on two integrated pillars: (1) the system of SBRs (Statistic Business Registers) which releases a variety of census-type data on the structural characteristics of units on an annual basis; and (2) the triennial Multi-Purpose (280,000 unit sample) Survey (MPS) which complements the information with additional data, primarily of a qualitative nature, on a range of topics (Istat, 2025b).

Istat, aware of certain limitations in European business statistics, developed the MPS to fill specific information gaps in the competitiveness of the business ecosystem and try to measure a complex and relevant phenomenon: the relationships between economic performance, organizational structures, and business strategies. To achieve this, quantitative data, administrative sources, or traditional classification variables (such as economic sector, enterprise size, and geographic location) are not enough. What is needed is the identification of behavioral profiles of firms, which can be captured through a multidimensional approach and qualitative questionnaires. Understanding firms’ strategic choices makes it possible to grasp the key drivers of the competitive evolution of the productive system and to account for its heterogeneity. The survey try to contribute

to shed some light on what economists refer as the “black box” of strategic behaviors or implicit competitiveness factors internal to firms and some recent economic analyses demonstrate this (Costa *et al.*, 2024; De Panizza *et al.*, 2024).

The questionnaire is divided into nine sections: Governance, ownership, control and management; Human resources; Production relation and supply chains; Market; Innovation and digitalization; Finance; Internationalization and global value chain; Smart Specialisation Strategy (S3); Environmental and social sustainability.

Below are two examples of innovative topics addressed in the MPS.

### ***3.1 New dimensions of analysis: the production “filières”***

The idea is to measure the economic dimension of the firm beyond its physical dimensions. The competitiveness of firms in a sector cannot be fully understood without considering what also happens to companies in the sectors upstream and downstream of the supply chains. For a long time, stakeholders and policymakers have called for robust information on the structure of production filières; accordingly, in close consultation with them, in the second edition of MPS a specific section was included in the questionnaire. The national data, with reference year 2022, were released in November 2023, the regional data at the end of April 2024. The Istat production filières have been widely used in the design of the Single Special Economic Zone (SEZ) by the Government. The Strategic Plan of the Single SEZ made by the Ministry for European Affairs, the South, Cohesion Policies and the NRRP approved with DPR (31/10/2025), identifies the strategic production filières to be strengthened, enhancing regional production vocations starting from those identified by Istat (Ministero delle Imprese e del Made in Italy, 2024).

### ***3.2 Environmental sustainability and economic performance***

The questionnaire adopted in the first edition has been further improved in the second one and allows to obtain an in-depth picture of the behaviors of firms in relation to environmental and social sustainability. This information combined with other variables from MPS and from registers allow to analyse also the relationship between environmental sustainability and economic performance.

According to the MPS 2022 almost 38.0% of Italian companies with three or more employees have carried out at least one environmental protection initiative. The percentage of firms that have invested in environmental sustainability increases from 34.5% of those with 3-9 employees to 73.8% of larger ones (250 or more employees). The propensity for environmental sustainability is significantly higher in industry than in services, with percentages ranging from 43.6% in micro-enterprises to 89.6% in large ones. A recent analysis (Istat, 2025c) focused on the manufacturing sector and concerns companies with at least 10 employees active in 2022 which employ approximately 2.9

million workers generating 266 billion in added value, equal to about one third of the companies, of the workers and of the national added value. These companies represent 18.1% of the units but 78.0% of the employment and 89.6% of the value added compared to the entire manufacturing sector. For those firms has been detected a positive and non-linear relationship between labour productivity and attention to sustainability: in particular, an increase in the environmental sustainability index corresponds to a “productivity bonus” that varies between 5% and 8%.

#### **4. The slow transition toward a greater use of new and alternative sources**

The challenge faced by NSIs is to remain relevant increasing the supply of information in line with the growing demand of users, despite declining real-term funding and growing competition in the business data arena (Menghinello *et al.*, 2020). At the same time, NSIs need to preserve the quality, transparency, and methodological rigor that official statistics must ensure, as well as ensuring a strong commitment to data confidentiality.

In recent years, there has been a growing trend toward reduced cooperation among respondents in business surveys, with response rates steadily declining over time (Erikson *et al.*, 2023). At national level, the obligation to respond (and related sanctions), provided for by art. 7 of DL n. 322/1989, has, until recently, allowed for maintaining adequate coverage but the problem must be addressed before it could seriously impact the quality of the results. Closely connected to business collaboration is the issue of the overall statistical burden, to which it is necessary to give concrete answers as part of a more general problem of administrative burden that the European Commission has committed to reduce (EU, 2025)

On the other hand, there are also well-known opportunities due to the growing digitalisation of all aspects of business and consequently the potential availability of new sources (big data, administrative data) and the new technologies to deal with them. But while, more and more alternative sources are emerging, the transition to the massive use of these sources for the NSIs is much more time demanding and resource consuming than expected, due to a complexity of institutional, cultural, organisational, practical as well as technical or methodological factors. A significant example of these difficulties is the recent case of the British Office for National Statistics (ONS) which, although made a formal strategic choice for a significant transition to alternative sources is still facing very serious problems. In the Report by the Office for Statistics Regulation/UK Statistical Authority release in April 2025 there is an entire section entitled “Making progress with administrative data is difficult” (Office for Statistics Regulation, 2025):

*“There have been some good examples of the use of administrative and big data sources in core economic statistics (...) However, as a whole, progress in using admi-*

*nistrative data from other government departments has been slow, reflecting in part practical and cultural challenges across government (...) these issues around data sharing are systemic and notes that there continues to be a failure to deliver on data sharing and linkage across government, alongside many persisting barriers to progress (...) Slow progress in the use of administrative data is particularly problematic in the context of reducing survey response rates. While noting that many of barriers may lie with the suppliers of data in other government departments, ONS staff observed that there was common acceptance that the use of administrative data should increase but no systematic assessment of the role that these data could play across output areas and the implications for the future role of surveys”.*

These issues should be complemented by the considerations written by Sir Robert Devereux in the Independent Review released in June 2025 (UK Government 2025) about ONS priorities for statistical production, a particularly delicate matter that, amid pervasive resource constraints, concerns also other NSIs.

*“There has been a commendable interest in both new approaches to statistics (including the use of administrative data), and ensuring the relevance of ONS activity to wider political debate. Unfortunately, this has had the (unintended) effect of de-prioritising the less exciting, but nonetheless crucial, task of delivering core economic statistics of sufficient quality to guide decision making”.*

These shortcomings have culminated in well-documented failures in economic statistics at the ONS. This experience provides a valuable lesson for all NSIs: while innovation and experimentation in accessing and leveraging new data sources are essential, it is equally important to define clear priorities in resource allocation, avoiding an excessive dispersion of efforts across multiple research streams that may produce numerous preliminary outputs and prototypes with limited industrial applicability, while diverting resources from other core statistics and potentially affecting their quality.

Also other countries found obstacles in accessing new sources and generally progress in those areas have been slower than expected (van Delden *et al.*, 2023; Eurostat, 2024). While the number of scientific publications and papers produced on these topics in the last 15 years has exploded, the same cannot be said about the concrete implementation in production by the NSIs, especially in key statistical domains, i.e. those in which European or national legal obligations are more stringent. In fact, still today in many countries business surveys continue to be the dominant source of data for official business statistics (Bender *et al.*, 2022; James *et al.*, 2020).

From this point of view, Italy continues to have some advantages compared to other countries. As said in section 2, Istat for years has accessed numerous alternative sources, especially administrative ones, and uses them pervasively in business statistics, often integrating them with direct surveys, creating a large register-based system ma-

king a lot of information available to users. This does not mean that there have not been, or still are, difficulties in accessing administrative data or significant delays in bringing key sources into regular production.

Since 2013, Istat has developed a line of research on the use of business websites as a Big Data source for official statistics, notably for updating and enriching the Istat Statistical Business Register and for supporting the survey on ICT usage in firms. In the latter case, the objective was to produce, through data-science approaches, a subset of the estimates currently obtained from questionnaires (e-commerce, etc.), while reducing response burden and, at the same time, generating new variables and greater sectoral and territorial detail with comparable quality. This work resulted in two release of experimental statistics on firms' use of websites (e.g. web ordering, use of social media, etc.), based on web scraping, text mining and machine-learning techniques and integrated with ICT survey data (Istat 2018; Istat 2020a). Despite their strong analytical and pioneering value in the scientific literature these experiences have not led to the regular production of current official statistics.

More recent and promising are studies on the application of web intelligence techniques for the automatic coding of products produced by enterprises in economic surveys (D'amore *et al*, 2025). The aim is to improve the quality and efficiency of surveys that use the product as the unit of analysis, for example the Annual Industrial Production Survey (PRODCOM), by automating the coding of products reported in questionnaires, supporting respondents during data collection, and substantially reducing the need for manual review. With the experience accumulated over the years and the recent developments in AI, this line of research will certainly lead to concrete results applicable in current production processes in a short time.

Like any other NSIs, Istat needs a clear internal strategy and operational roadmap, multidisciplinary skills and appropriate technical infrastructure to address the complex challenges related to administrative and big data, together with strong inter-institutional cooperation, an effective data stewardship role and international collaboration for the exchange of best practices.

Within this challenging context, the strategic approach that aims at enhancing information collection while reducing the burden on firms and costs to the statistical system, calls for the broad adoption of clear principles across the NSIs to be implemented through a balanced and flexible framework, as outlined below.

- Exploit all currently existing data before collecting additional data; see, for example, Statistic Canada's "Admin-First" strategy (Rancourt, 2018); make also extensive use of the possibility of microdata linking, for example, SBS with ICT (Istat, 2020b) or trade statistics with gender statistics (Istat-Ice, 2025) to produce new statistics.

- Rationalize direct surveys and, where possible, optimize and reduce questionnaires.
- Adopt criteria inspired by reuse “by design” so that all the tools and data produced can be easily integrated/adapted to meet new needs.
- Prioritize resource allocation and focus on high-impact trajectories, data sources, and innovations enabling rapid transition to operational deployment.
- Exploit new “high-value/high-impact” sources with economies of scale and scope.

### **5. An innovative and high-value source: electronic invoices and the application to short term statistics**

In Italy, over the last five years, electronic invoicing has been progressively introduced for all economic operators. It is mandatory for business-to-business (B2B), business-to-government (B2G) and business-to-consumer (B2C) transactions and related invoices are processed electronically within the Exchange System under the supervision of the Revenue Agency (RA). From a statistical point of view, this new tax source offers unique advantages: complete coverage of firms, timeliness, and the presence of relevant information including the link between seller and buyer. This source can bring about a substantial transformation in economic statistics, significantly improving their timeliness, quality and at the same time reducing the administrative burden for firms, especially in short-term statistics. But its potential uses in economic statistics and national accounts are countless (business registers updated in real time, inter-sectoral tables; estimation of goods and services flows between suppliers and customers to map the network of inter-firm exchanges; etc.). It represents a revolution in economic statistics.

Between 2020 and 2023, Istat conducted experimental activities (Rapiti *et al.* 2023) on a sample of firms to assess microdata quality, finalizing a framework agreement with the Revenue Agency (RA) in June 2024. Since summer 2024, Istat has been receiving the full population of e-invoice microdata on a monthly basis, with a 29-day delay relative to the reference month (about 135 million records per month).

Although structured, this source is comparable in scale to big data. Each transaction includes information on the seller, buyer, turnover, issue date, and type of counterparty. For privacy reasons, buyers are identified only for B2B transactions, and invoice contents (products and prices) are not provided.

Since summer 2025, Istat has also begun receiving monthly B2C e-receipt data (e.g. restaurants, hotels, and retailers). The integration of these sources supports a multi-year strategy to progressively integrate and replace survey data, redesigning turnover and retail sales surveys, reducing their size, limiting manual data revision, and significantly

lowering the reporting burden on firms. The reorganization of the short-term surveys will allow produce better estimates based on the entire population of firms and to reduce the statistical burden: in a short time, this approach is expected to relieve firms of approximately 450,000 monthly reporting obligations every year. This marks only the beginning of a long process of systematic use of e-invoicing data, which is expected to deliver the most significant reduction in administrative burden ever achieved at the national level.

Up to 2025, the use of e-invoice data has been limited to imputing missing monthly responses in the industry and services surveys. The multi-year plan envisages a gradual, three-phase replacement of survey data with administrative sources starting in January 2026. First, in the services survey (currently based on a 25,000-unit sample), turnover estimates for 11,000 small and medium-sized enterprises (SMEs) will be replaced with e-invoices data (-40% of the sample units), while moving from a sampling approach to using the full population of companies available. Second, still for services, this substitution will be extended to accommodation and food-service activities using e-receipts data; in parallel, in the industry survey (now covering 8,000 firms), e-invoice will replace survey data for most SMEs, and the majority of units in the monthly retail trade survey will also be taken over by administrative sources. Third, further steps may include discontinuing some direct surveys; however, for very large enterprises – at least in industry – direct data collection is likely to remain necessary for methodological and data-quality reasons.

The extensive use of e-invoice and e-receipt data to produce and disseminate highly timely monthly statistics entails certain risks, stemming from Istat's full dependence on an external provider that may not always guarantee timely delivery for technical or administrative reasons, as well as from residual quality issues in the source data. For this reason, a prudent approach is advisable, favouring a gradual transition from surveys to administrative data and the preparation of a contingency plan. (Erikson *et al.*, 2023).

## 6. Conclusions

In a world with rapidly changing data and information environment, and with high uncertainty in the global economy, NSIs have the challenging role to provide critical information helping policy makers and the public make more informed decisions in troubled waters. In this context, up-to-date and future-proof business and trade statistics play a key role.

In Italy, as previously illustrated, the system for producing and disseminating business and trade statistics is both modern and effective. In recent years, numerous innovations and developments have taken place, driven by the EBS Regulation and linked to priorities defined at national level.

The current challenge is to produce and disseminate new and existing business statistics in innovative, smarter and less burdensome ways achieving a tangible reduction in the burden on firms, wherever feasible. The path undertaken must therefore be continually improved through the use and potential integration of surveys with new sources (UNECE, 2018; de Waan *et al.*, 2020). Literature has long made it clear that new data sources and direct surveys are increasingly complementary tools (Snijkers *et al.*, 2023; van Delden *et al.*, 2023) but emphasis should be shifted towards the former, progressively restricting the latter only where strictly warranted. Keeping in mind that very often less burden and work on business put more work on ourselves as official statisticians. Below are some insights and key action points derived from Istat concrete experience in recent years.

1. Continued exploration and exploitation of new administrative, commercial or big data sources, including privately held data, for “quantitative” information, increasingly combined with innovative methods (e.g., artificial intelligence, web intelligence, etc.);
2. A corresponding reduction in direct business surveys for quantitative data collection should be promoted, without loss of informational content, as these are replaced by the sources cited in point 1 above.
3. Maintain a balanced and non-excessive use of (mainly) qualitative direct surveys (e.g. MPS in Italy) to collect information not available from alternative sources—particularly regarding business strategies and emerging topics.
4. Removal of remaining institutional, organizational, cultural, and practical barriers to the timely access of administrative data (OECD, 2025) A more enabling and cooperative environment must be established to meet statistical needs, which is a national responsibility and must be fulfilled through coordinated efforts by all institutional actors. National and EU legislation should further concretely and effectively strengthen the role of statistical authorities in accessing administrative data.
5. Increased inter-institutional cooperation (including microdata sharing) with other data-producer to avoid duplication in data collection and to enhance consistency (see the collaboration between Istat and the Bank of Italy on STEC);
6. Expansion of microdata sharing initiatives among EU countries (such as MDE), and removal of existing obstacles to facilitate broader exchange;
7. The massive use of firm microdata integration among different sources increases data privacy risks (especially for sole proprietorship firms). Therefore always ensure compliance with current national and European (GDPR) legislation regarding the protection of personal data.

## References

- BENDER S., SAKSHAUG J. W. (2022). Data Sources for Business Statistics: What has Changed?. *The Survey Statistician*, 10-18 Vol 85
- COSTA S., DE SANTIS S., DOSI G., SBARDELLA A., VIRGILLITO M. E. (2024). Ristrutturazioni del tessuto produttivo tra pandemia e inflazione: comportamenti, persistenze e transizioni d'impresa. *Economia Italiana* n. 1
- D'AMORE G., MORRONE M., SUMMA D. (2025). Application of web intelligence techniques for the automatic coding of products manufactured by businesses in economic surveys. Presented at *UNECE Expert Meeting on Statistical Data Collection and Sources* 2-4 June 2025, Lisbon, Portugal
- DE PANIZZA A., FARAMONDI A., LOMBARDI S., TRUGLIA F. G. (2024). The Italian production system in the aftermath of the pandemic: firms' behaviours, production filières and productivity. *Economia Italiana* n. 1
- DE WAAL, T., VAN DELDEN, A., & SCHOLTUS, S. (2020). Multi-source statistics: Basic situations and methods. *International Statistical Review*, 88(1), 203-228. <https://doi.org/10.1111/insr.12352>
- ERIKSON J., GIESEN D., HOUBEN L., SARAIVA P. (2023). Managing Response Burden for Official Statistics Business Surveys – Experiences and Recent Developments at Statistics Netherlands, Statistics Portugal, and Statistics Sweden. In SNIJKERS G., BAVDAŽ M., BENDER S., JONES J., MACFEELY S., SAKSHAUG J., THOMPSON K. J., VAN DELDEN A. (Eds.), *Advances in Business Statistics, Methods and Data Collection*, Wiley
- EU (2025). Commission proposes to cut red tape and simplify business environment. Communication. Retrieved from: [https://commission.europa.eu/news-and-media/news/commission-proposes-cut-red-tape-and-simplify-business-environment-2025-02-26\\_en](https://commission.europa.eu/news-and-media/news/commission-proposes-cut-red-tape-and-simplify-business-environment-2025-02-26_en)
- EUROPEAN PARLIAMENT AND THE COUNCIL (2019). Regulation (EU) 2019/2152 of the of 27 November 2019 on European business statistics, (known as the EBS Regulation)
- EUROSTAT (2017). European Statistics Code of Practice — revised edition 2017. Retrieved from: <https://ec.europa.eu/eurostat/web/products-catalogues/-/ks-02-18-142>
- EUROSTAT (2024). Result of Third round of peer reviews (2021-2023). Retrieved from: <https://ec.europa.eu/eurostat/web/quality/peer-reviews/third-round>
- JAMES G., McLAREN C., (2020). Measuring the Economy Appendix Sources for economic statistics, Retrieved from: <https://www.measuringtheeconomy.uk/book/text/50-01-appendix-01.html>
- ISTAT (2016). Quality Guidelines for statistical processes using administrative sources Retrieved from: <https://www.istat.it/wp-content/uploads/2023/12/Linee-Guida-v1.1-Versione-inglese.pdf>
- ISTAT (2018). Modalità di utilizzo dei siti web da parte delle imprese. Statistica sperimentale. Retrieved from: <https://www.istat.it/statistica-sperimentale/stime-delle-modalita-di-utiliz>

zo-dei-siti-web-da-parte-delle-imprese/

ISTAT (2020a). Le distribuzioni dei siti web delle imprese secondo i servizi e le funzionalità offerte: il processo di stima supportato dall'uso di dati da Internet. Statistica sperimentale. Retrieved from: [https://www.istat.it/wp-content/uploads/2020/05/Nota\\_Metodologica.pdf](https://www.istat.it/wp-content/uploads/2020/05/Nota_Metodologica.pdf)

ISTAT (2020b). Enterprises classified by use of ICT and economic indicators –Methodological Note. Retrieved from: <https://www.istat.it/en/experimental-statistic/integration-between-data-from-register-and-sample-surveys-enterprises-classified-by-use-of-ict-and-economic-indicators/>

ISTAT (2024a). Turnover of industry and services - January 2024 Press Release. Retrieved from: <https://www.istat.it/en/press-release/turnover-of-industry-and-services-january-2024/>

ISTAT (2024b). Services producer prices - Q1 2024 Press Release Retrieved from: <https://www.istat.it/en/press-release/services-producer-prices-q1-2024/>

ISTAT (2025a). Foreign-trade-and-Import-prices Press Release. Methodological Note Istat. Retrieved from: [https://www.istat.it/wp-content/uploads/2025/06/Foreign-trade-and-Import-prices\\_April2025.pdf](https://www.istat.it/wp-content/uploads/2025/06/Foreign-trade-and-Import-prices_April2025.pdf)

ISTAT (2025b). Piano generale del Censimento permanente delle imprese. Retrieved from: <https://www.istat.it/wp-content/uploads/2025/09/Piano-Generale-Censimento-2025.pdf>

ISTAT (2025c). Sostenibilità ambientale e performance economica delle imprese manifatturiere, anno 2022. Statistica Focus. Retrieved from: <https://www.istat.it/comunicato-stampa/sostenibilita-ambientale-e-performance-economica-delle-imprese-manifatturiere-anno-2022/>

ISTAT-ICE (2025). Annuario statistico Commercio estero e attività internazionali delle imprese Edizione 2025. Retrieved from: <https://annuarioistatice.istat.it/avvio.html>

MENGHINELLO S., FARAMONDI A., LAURETI T. (2020). The future role of official statistics in the business data arena. *Statistical Journal of the IAOS*. Volume 36, Issue 2. <https://doi.org/10.3233/SJI-190557>

MINISTERO DELLE IMPRESE E DEL MADE IN ITALY (2024). Piano strategico della ZES unica, Retrieved from: <https://www.strutturazes.gov.it/media/itmfe/trr/piano-strategico-zes-integrale.pdf>

OECD (2025). Recommendation on Enhancing Access to and Sharing of Data. Retrieved from: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463>

OFFICE FOR STATISTICS REGULATION (2025). Systemic Review of ONS Economic Statistics Interim Report. April. Retrieved from: <https://osr.statisticsauthority.gov.uk/publication/systemic-review-of-ons-economic-statistics-interim-report/pages/8/also%20https://osr.statisticsauthority.gov.uk/publication/data-sharing-and-linkage-for-the-public-good-follow-up-report/pages/3/>

RANCOURT E. (2018). Admin-First as a Statistical Paradigm for Canadian Official Statistics: Meaning, Challenges and Opportunities. *Proceedings of Statistics Canada Symposium* Retrieved from: <https://www.statcan.gc.ca/en/media/2295>

RAPITI F. M., FILIBERTI S., BRUNO G., AMATO G., ORSINI D. (2023). E-invoices as a new timely ad-

ministrative (and big) data source for short term business statistics. Presented at *Conference EESW23*, Lisbon. Retrieved from: <https://drive.google.com/file/d/18oNcYXm5Ea-Az0Lr-czFgIOKfaugsHWlm/view>

SNIKERS G., BAVDAŽ M., BENDER S., JONES J., MACFEELY S., SAKSHAUG J., THOMPSON K. J., VAN DELDEN A. (2023). *Advances in Business Statistics, Methods and Data Collection: Introduction*. In SNIKERS G., BAVDAŽ M., BENDER S., JONES J., MACFEELY S., SAKSHAUG J., THOMPSON K. J., VAN DELDEN A. (Eds.), *Advances in Business Statistics, Methods and Data Collection*, Wiley

UNECE (2018). *A Guide to Data Integration for Official Statistics (Version 2.0)*. Retrieved from: <https://unece.org/sites/default/files/2024-07/HLG-MOS%20Guide%20to%20Data%20Integration%20for%20Official%20Statistics.pdf>

VAN DELDEN A., LEWIS D. (2023). *Methodology for the Use of Administrative Data in Business Statistics* in In Snijkers G., Bavdaž M., Bender S., Jones J., MacFeely S., Sakshaug J., Thompson K. J., van Delden A. (Eds.), *Advances in Business Statistics, Methods and Data Collection*, Wiley

UK GOVERNMENT (2025). *Independent Review by Sir Robert Devereux KCB*. Published 26 June 2025. Retrieved from: <https://www.gov.uk/government/publications/independent-review-of-the-performance-and-culture-of-the-office-for-national-statistics/independent-review-by-sir-robert-devereux-kcb-june-2025>

## LA RIDUZIONE DELLA DIMENSIONALITÀ DI UN SET DI INDICATORI CONGIUNTURALI PER LA STIMA TRIMESTRALE DEL VALORE AGGIUNTO DEI SERVIZI NEI CONTI ECONOMICI TRIMESTRALI: METODOLOGIE A CONFRONTO

### REDUCING THE DIMENSIONALITY OF A SET OF SHORT TERM INDICATORS FOR THE QUARTERLY ESTIMATE OF THE VALUE ADDED IN SERVICES SECTOR IN THE QUARTERLY NATIONAL ACCOUNTS: A COMPARISON OF METHODOLOGIES

*Marianna Ascione<sup>1</sup>*

#### Sommario

Il lavoro<sup>2</sup> descrive alcune innovazioni introdotte nella stima trimestrale del valore aggiunto per il settore dei servizi che confluisce nei Conti economici trimestrali (CET) – in termini di nuove fonti e nuove metodologie nell’ultima revisione straordinaria programmata dall’Istat. Al fine di migliorare la rappresentazione di questo settore economico è stato deciso di adottare uno schema di classificazione delle attività economiche (ATECO) più esteso per le fasi di elaborazione intermedia.

Per la stima dei CET non essendo disponibile informazione diretta ovvero l’osservazione del fenomeno al livello trimestrale, si ricorre a tecniche econometriche di disaggregazione temporale a partire da indicatori congiunturali la cui dinamica verosimilmente descrive l’informazione disponibile soltanto a livello annuale. E’ stato pertanto necessario passare al vaglio centinaia di nuovi indicatori congiunturali per verificare la possibilità di adottarne alcuni come input delle tecniche di disaggregazione temporale per taluni settori. Il lavoro descrive quanto implementato per il settore del commercio al dettaglio.

Per tale scopo sono state messe a confronto tecniche di *variable selection (Shrinkage Methods & Methods using derived directions)* al fine di individuare gli indicatori più opportuni da utilizzare come predittori ottimali nella stima di uno dei più rilevanti dei settori dei servizi.

---

<sup>1</sup> Istat - Istituto Nazionale di Statistica, Roma, Italia - email: ascione@istat.it

Le opinioni contenute in questo contributo sono espresse dall’Autrice a titolo personale e non coinvolgono l’Istituzione di appartenenza.

<sup>2</sup> Il presente lavoro costituisce una sintesi della prova finale del Master Universitario di II livello in “Big Data. Metodi statistici per la società della conoscenza” presso l’Università di Roma, Sapienza - Dipartimento Scienze Statistiche. I risultati raggiunti sono stati implementati nei CET in occasione della revisione straordinaria dei Conti economici rilasciata nel 2019. L’approccio è stato confermato nella revisione straordinaria successiva programmata a fine 2024.

## **Abstract**

*The work describes some innovations introduced in the quarterly estimate of value added for the services sector that is included into the Quarterly national accounts (QNA) - in terms of new sources and new methodologies during the extraordinary revision scheduled by Istat. In order to improve the representation of this economic sector, the decision to adopt a more extensive classification scheme for economic activities (ATECO) for the intermediate processing phases was made.*

*For the QNA estimate, since direct information namely the observation of the phenomenon at the quarterly level is not available, econometric techniques of temporal disaggregation are used starting from short-term indicators whose dynamics likely describes the information available only at an annual level. It was therefore necessary to screen hundreds of new economic indicators to verify the possibility of adopting some of them as input of the temporal disaggregation techniques for certain sectors. The work describes what has been implemented for the retail sector. For this purpose, variable selection techniques (Shrinkage Methods & Methods using derived directions) were compared in order to identify the most appropriate indicators to be used as optimal predictors in estimating one of the most relevant services sector.*

**Keywords:** *Big data, Variable reduction, Penalized regression, Principal component analysis, Quarterly national accounts.*

## **1. Introduzione**

A differenza della contabilità annuale, ad alta frequenza non si dispone della stessa quantità di informazioni (serie mensili/trimestrali), pertanto nell'ambito dei Conti economici trimestrali<sup>3</sup> (CET) si ricorre a tecniche econometriche (metodo indiretto) con cui si stima la dinamica infrannuale di un fenomeno economico tramite un'altra variabile (*proxy*) fortemente correlata alla prima. Si veda di Chow – Lin (1971). È necessario quindi valutare la relazione esistente tra i valori annuali di una variabile economica e quelli annualizzati di un indicatore congiunturale di riferimento che è invece disponibile su base infrannuale. Tali tecniche econometriche utilizzano la relazione descritta per operare la disaggregazione temporale della serie disponibile soltanto a bassa frequenza (serie annuali).

In corso d'anno, ovvero quando non si dispone del dato annuale più recente, l'informazione trimestrale è estrapolata in funzione della relazione indicato-indicatore che è stata stimata sull'intero periodo.

<sup>3</sup> Si veda la nota informativa I Conti economici trimestrali. Principali elementi informativi rilasciata dall'Istat a settembre 2015, che descrive il metodo indiretto di trimestralizzazione e l'elenco dei principali indicatori congiunturali utilizzati nelle stime. La nota è stata aggiornata a ottobre 2024.

Per la compilazione dei CET si utilizzano svariati indicatori congiunturali in fase di trimestralizzazione che sono tratti da diverse fonti Istat o esterne all'istituto (ad esempio per la stima del valore aggiunto del settore industria, si utilizzano gli indici mensili della produzione industriale e del fatturato dell'industria). Si veda Istat (2015) e Istat (2024).

La revisione<sup>4</sup> dei conti economici effettuata nel 2019 è stata l'occasione per ricercare nuovi indicatori congiunturali per verificarne la performance previsiva e confrontarli con gli indicatori correntemente utilizzati nelle procedure di trimestralizzazione nei vari domini.

Nel presente lavoro più che porre l'enfasi sulle tecniche di disaggregazione temporale – peculiari per la contabilità trimestrale, vengono descritte le scelte operate per ricercare gli indicatori congiunturali più idonei a descrivere la dinamica infrannuale del commercio al dettaglio alla luce della scarsità di informazioni disponibili per la rappresentazione del settore. L'impianto metodologico proposto è stato confermato anche nell'ambito della recente revisione dei conti economici che ha portato alla diffusione di serie riviste a fine 2024.

Il lavoro è strutturato come segue: nella sezione 2 si descrive il problema relativo alla stima trimestrale in generale e per i servizi, in particolare; la sezione 3 è dedicata alla descrizione delle fonti informative da cui si è partiti per la costruzione di un ampio set di serie storiche con differente cadenza temporale; la sezione 4 descrive i diversi approcci metodologici utilizzati per la riduzione della dimensionalità dei dati e per la costruzione di indicatori da utilizzare in fase di disaggregazione temporale; i risultati empirici sono riportati nella sezione 5; le considerazioni conclusive sono riportate nella sezione 6. Il lavoro nella sua forma più estesa, infine, riporta la descrizione degli approcci metodologici implementati in appendice.

## **2. Descrizione del problema: la stima trimestrale dei servizi nei Conti economici trimestrali**

Nei Conti economici trimestrali per la stima del valore aggiunto dei servizi si utilizzano le serie rese disponibili dall'indagine Istat sul fatturato dei servizi. Tali indici costituiscono la *proxy* impiegata nella procedura di trimestralizzazione delle serie annuali del corrispondente aggregato di contabilità nazionale.

---

<sup>4</sup> Per revisione straordinaria dei conti (benchmark) si intende l'insieme di operazioni effettuata generalmente a cadenza quinquennale ed in corrispondenza degli anni con cifra finale 4 e 9 per incorporare le modifiche concordate con Eurostat (ovvero applicare l'insieme di regole di compilazione definite nella nuova versione del Sistema Europeo dei Conti, SEC2010) e le innovazioni che scaturiscono dalla disponibilità di nuove fonti e da eventuali innovazioni nelle metodologie di misurazione dei fenomeni economici rappresentati.

Il settore dei servizi – il cui valore aggiunto confluisce nella stima del Pil – include anche il commercio. Il presente lavoro descrive il contributo al miglioramento della stima trimestrale del commercio e di conseguenza del Pil trimestrale rilasciato a t+60 giorni dal trimestre di riferimento e della stima preliminare rilasciata a t+30 dal trimestre di riferimento.

A partire dal 2019 il settore viene stimato distintamente per le seguenti tipologie di attività: commercio e manutenzione di autoveicoli (ATECO G45), commercio all'ingrosso (G46) e commercio al dettaglio (G47). Prima della revisione, le ATECO G45 E G47 erano trattate in maniera aggregata.

L'indagine sul fatturato dei servizi, divenuta nel frattempo disponibile a cadenza mensile, fornisce un set completo di indicatori per la stima soltanto per il rilascio del Pil a t+60 giorni dal trimestre di riferimento. Tali indicatori sono utilizzati per la stima trimestrale del valore aggiunto del commercio e manutenzione di autoveicoli e del commercio all'ingrosso. Da regolamento tale indagine non copre il settore G47 che invece viene rappresentato attraverso l'indice del commercio al dettaglio elaborato dall'Istat a cadenza mensile.

Il passaggio allo schema di classificazione a 45 branche di attività economiche adottato per esigenze di elaborazione intermedia e per aumentare la granularità dei dati, ha fatto emergere la necessità di cercare nuovi indicatori da utilizzare quale proxy nelle procedure di disaggregazione temporale per la branca 23 corrispondente al commercio al dettaglio (ATECO G47). In teoria l'indicatore ideale alla rappresentazione del valore aggiunto del settore relativo al commercio al dettaglio è costituita dagli indici elaborati a partire dai dati raccolti con la rilevazione mensile sulle vendite al dettaglio condotta dall'Istat. Tale indagine si riferisce alle imprese commerciali che svolgono come attività prevalente la vendita al dettaglio (come definita nella Classificazione ATECO); tuttavia la correlazione tra la serie annualizzata dell'indice e la serie annuale della contabilità annuale non risulta soddisfacente .

La carenza di opportuni indicatori congiunturali è ancora più avvertita per la stima anticipata del Pil che viene rilasciata a t+30 giorni dal trimestre di riferimento in quanto le fonti informative hanno una copertura ancora più bassa non essendo disponibili in tempo per le esigenze di elaborazione.

Per far fronte a tale esigenza informativa è stata effettuata una ricognizione delle fonti informative disponibili ad alta frequenza (mensili e trimestrali) descritte nella sezione che segue.

### 3. Fonti informative

L'ipotesi alla base di questo contributo è che il commercio al dettaglio sia condizionato da svariati fattori; sono stati pertanto considerati centinaia di indicatori a cadenza mensile e trimestrale che descrivono numerosi fenomeni economici. Questa forma di commercio<sup>5</sup> inoltre è influenzata anche dal fatto che i commercianti al dettaglio possono vendere tanto ad altre branche di attività economica quanto alle famiglie in veste di consumatori finali.

Poiché la principale fonte disponibile relativa alle vendite al dettaglio risulta scarsamente correlata con la variabile annuale di contabilità nazionale e in assenza di altra informazione diretta in grado di rappresentare il settore, è stato censito l'insieme di indicatori congiunturali prodotti dall'ISTAT, dalla Banca d'Italia sul Sistema dei pagamenti e di un altro set di indicatori derivato sui margini commerciali frutto di elaborazioni intermedie a partire dai dati di contabilità nazionale.

La tavola 1 di seguito riportata, riassume l'insieme di serie considerate nell'analisi. Come si può notare dalla tavola, sono utilizzati dati provenienti da vari domini. L'eterogeneità delle fonti pertanto assicura la "varietà" che è una delle caratteristiche che inducono a definire *big data* le tipologie di dati utilizzati.

Si veda IMF (2017) e Eurostat (2018).

---

<sup>5</sup> Il commercio al dettaglio è classificato in primo luogo in base al tipo di esercizio nel quale viene effettuata la vendita (ad esempio, in negozio o al di fuori del negozio), al grado di specializzazione, al tipo di prodotti venduti. Secondo la classificazione statistica delle attività economiche (NACE Rev.2) questa divisione include la rivendita (vendita senza trasformazione) di beni nuovi e usati destinati principalmente ad uso e consumo personale o domestico in negozi, grandi magazzini, mercati e banchi, attraverso imprese di vendita per corrispondenza, mediante venditori porta a porta, venditori ambulanti, cooperative di consumo, via internet eccetera.

*Tavola 1. Indicatori congiunturali utilizzati nell'esercizio per il periodo 1995Q1-2017Q4*

<b>Dominio</b>	<b>Numero serie</b>	<b>Cadenza</b>	<b>Fonte</b>
Produzione industriale per prodotto	40	M	Istat
Fatturato dei servizi	25	M	Istat
Fatturato e ordinativi dell'industria	69	M	Istat
Commercio estero	100	M	Istat
Serie elaborate sui margini commerciali e del trasporto	171	Q	Istat
Vendite al dettaglio	20	M	Istat
Forze lavoro, occupazione ed ore lavorate	32	M/Q	Istat
Clima di fiducia delle imprese dei servizi	40	M	Istat
Spese per consumi delle famiglie	63	Q	Istat
Dati sulle immatricolazioni	17	M	Unrae
Dati sui premi delle assicurazioni	2	Q	Ivass
Serie di natura finanziaria (sifim, tassi di interesse, depositi e prestiti)	52	Q	Banca d'Italia
Serie sul Sistema dei pagamenti	36	M/Q	Banca d'Italia

M = serie mensile; Q = serie trimestrale; M/Q = serie a cadenza sia mensile sia trimestrale nel dominio.

Fonte: elaborazione propria su fonti disponibili per le finalità dei CET

Una volta raccolte le informazioni dai vari domini, l'intero set di dati è stato oggetto di un pre-trattamento come descritto nel paragrafo che segue.

#### **4. Principali passi di implementazione delle tecniche di riduzione della dimensionalità e costruzione di indicatori sintetici**

L'analisi della relazione indicato/indicatore in questo esercizio è stata effettuata a partire dai dati annuali per il periodo 2004-2017 (14 osservazioni per 667 regressori).

Per quanto riguarda gli indicatori raccolti, per un numero limitato di domini è stata necessaria la ricostruzione delle serie che da regolamento comunitario, rendono disponibili i dati a partire da un certo anno di inizio successivo a quello necessario alla stima dei CET (1995Q1). Le serie relative a tutti gli indicatori congiunturali disponibili a diversa cadenza temporale sono stati dapprima aggregate per ricondurle a serie trimestrali

e successivamente standardizzate in quanto espresse in differenti unità di misura. La standardizzazione inoltre ha consentito il ricorso ai metodi implementati per lo shrinkage (*regularized regression*) che non sono invarianti rispetto a trasformazioni di scala. Tibshirani (1996).

In generale, il processo di elaborazione dei CET prevede per ciascuna branca di attività economica la stima trimestrale di produzione, consumi intermedi e valore aggiunto espresso in termini nominali, ai prezzi dell'anno precedente ed in volume. L'analisi di seguito descritta, è stata circoscritta alla sola serie espressa in volume; pertanto per le fasi successive di questo lavoro, il valore aggiunto annuale di contabilità nazionale nel settore del commercio al dettaglio, è da considerarsi come variabile dipendente dei modelli di seguito applicati. Per la descrizione del processo di stima per produzione e consumi intermedi si veda la nota metodologica Istat (2015) e Istat (2024).

Per la riduzione della dimensionalità dei dati sono stati considerati diversi approcci. Poiché il modello di regressione lineare (stime OLS) non garantisce una buona performance quando si dispone di un ampio set di dati multivariati contenente un numero di variabili superiori al numero delle osservazioni sono state considerate come una alternativa valida le tecniche di regressione penalizzata (noti come *shrinkage/regularization methods*) che aggiungono un vincolo nell'equazione di stima dei coefficienti. Valori dei coefficienti nulli o prossimi allo zero sono la conseguenza principale di tale restrizione.

Le tecniche di regressione penalizzata sono generalmente utilizzate in aggiunta ai consueti metodi di regressione basati su tecniche di riduzione dimensionale che pure risultano molto utili con set di dati di grandi dimensioni con più variabili predittive correlate. Nel presente lavoro è stato utilizzato il metodo lasso<sup>6</sup> che è indicato nei modelli che presentano sparsità nei dati ovvero in presenza di numerosi regressori non rilevanti e correlati tra di loro (multicollineari come quelli utilizzati). Tibshirani (1996).

I metodi di riduzione delle dimensioni che partono da una combinazione lineare dei predittori selezionati, restituiscono nuove variabili – generalmente in numero inferiore alle variabili di partenza – chiamate di seguito fattori o componenti principali – a seconda del metodo implementato – tra loro incorrelati che vengono utilizzati come predittori a loro volta ovvero come variabili esplicative del modello. Per la stima del valore ag-

<sup>6</sup> Lasso utilizza la norma L1 che corrisponde alla somma dei coefficienti in valore assoluto. Definita la somma dei quadrati dei residui  $RSS=(y - X\beta)'(y - X\beta)$  lo stimatore proposto dal metodo è:  $L=1tRSS+$ . Per la selezione con questo metodo, è necessario definire un parametro  $\lambda$  detto di *tuning/ottimizzazione* che determina la quantità di restrizione (*shrinkage*) ovvero l'impatto della penalità (*penalty*): con  $\lambda = 0$ , tale termine non ha alcun effetto e la regressione restituisce una stima dei coefficienti OLS; all'aumentare di  $\lambda$ , l'impatto aumenta ed i coefficienti di regressione si approssimano allo zero fino – in alcuni casi – ad annullarsi, Tibshirani (1996). L'implementazione per questo esercizio è avvenuta con il package *biglasso* sviluppato in R (Yaohui Z., Chuyi W., Tabitha P., Breheny P.).

giunto in volume del commercio al dettaglio, variabile dipendente del modello, è stata seguita una procedura suddivisa in più fasi di seguito descritta:

1. analisi preliminare della correlazione tra la variabile annuale di contabilità nazionale e ciascun indicatore congiunturale annualizzato utilizzando tale informazione per la tecnica ad-hoc;
2. selezione dei regressori mediante tecniche alternative di *variable selection*: un metodo di shrinkage (lasso regression come in Tibshirani (1996)) e una tecnica ad-hoc implementata sulla base di una elevata correlazione (almeno 0.9% per il periodo 2006-2016) tra indicato/indicatore e basse innovazioni in termini di errore medio assoluto (MAE) ottenute con il metodo di trimestralizzazione di Fernandez (1981);
3. sintesi dei regressori ottimali individuati al passo precedente mediante l'estrazione di componenti principali (ACP) come in Hotelling (1933) e Stock e Watson (2002) e di fattori con una analisi canonica (ACC) come in Box e Tiao (1977). Queste due differenti impostazioni hanno consentito di affrontare la riduzione del problema dimensionale attraverso sia un'impostazione statica della regressione (ACP) sia mediante una rappresentazione dinamica delle relazioni con le variabili indipendenti che hanno un effetto ritardato sulla variabile dipendente (ACC).
4. analisi comparativa della performance dei diversi modelli ottenuti attraverso l'implementazione di un modello di tipo  $ADL(1,1)$ <sup>7</sup> sulla base del MAE sui tassi di crescita annuali calcolati a partire dalle serie trimestrali stimate con i differenti regressori. Anche i modelli ADL utilizzati forniscono una rappresentazione dinamica della relazione descritta dal modello di regressione per il valore aggiunto.

## 5. Analisi dei risultati

L'analisi si basa su 2 set di indicatori: un primo set di 14 indicatori selezionati con il metodo *ad-hoc* ed un secondo set di 10 indicatori individuati con il metodo lasso entrambi da utilizzare come variabili esplicative nel modello costruito per la variabile dipendente valore aggiunto.

<sup>7</sup> Il modello a ritardi distribuiti  $ADL(1,1)$  correntemente utilizzato nei CET. Nella forma state-space sono utilizzati per disaggregare gli aggregati annuali di contabilità nazionale a livello trimestrale, utilizzando il filtro di Kalman in aggiunta ad un algoritmo di smoothing. Da questa rappresentazione si possono ricavare i modelli di Chow – Lin (1971) e di Fernandez (1981) per opportuni vincoli sui parametri. Il primo modello corrisponde ad una regressione statica con errori distribuiti come un  $AR(1)$  e risulta molto stabile nelle estrapolazioni; il secondo invece ad una regressione statica con disturbi con un andamento simile ad un processo random walk (RW). Entrambi garantiscono elevate correlazioni tra stime disaggregate e indicatori impiegati in termini di tassi di crescita. Si veda Proietti (2006), Bisio and Moauro (2018).

La tavola 2 di seguito riportata contiene gli indicatori selezionati con i due metodi ordinati in funzione della rilevanza (*loadings*) nella spiegazione della variabile dipendente (fase 2).

*Tavola 2. Indicatori congiunturali selezionati*

Serie selezionate con il metodo <i>ad-hoc</i>		Serie selezionate con il metodo lasso	
1	fatturato industria - componente estera - beni intermedi	1	importazioni extra UE per il prodotto 48 [altri servizi: nel campo dell'arte, dello spettacolo e del tempo libero]
2	importazioni intra UE totale	2	spese per consumi delle famiglie funzione [servizi ricreativi]
3	indice derivato sui margini del commercio	3	indice derivato sui margini di trasporto per il prodotto 14 [macchine ed apparecchi meccanici]
4	fatturato dei servizi - totale	4	esportazioni intra UE - totale
5	importazioni intra UE per il prodotto 13 [apparecchiature elettriche]	5	importazioni intra UE per il prodotto 16 [altri mezzi di trasporto]
6	importazioni extra UE per il prodotto 10 [articoli in gomma e materie plastiche]	6	produzione nel settore della pesca in volume
7	fatturato industria - componente estera - consumi totali	7	esportazioni intra UE per il prodotto 14 [macchine ed apparecchi meccanici]
8	spese per consumi delle famiglie per la funzione 7 [servizi per la casa (affitti, condominio, manutenzione, ecc.)]	8	esportazioni extra UE per il prodotto 17 [mobilio; altri manufatti (esclusi gioielli); servizi di riparazione e installazione di macchinari e apparecchi]
9	produzione industriale per il prodotto 15 [autoveicoli, rimorchi e semirimorchi]	9	esportazioni intra UE per il prodotto 1 [prodotti dell'agricoltura, della caccia e della silvicoltura]
10	spese per consumi delle famiglie in volume per la funzione 17 [combustibili]	10	fatturato industria - componente domestica (ATECO B)
11	importazioni extra UE per il prodotto 18 [gioielli e articoli di oreficeria]		
12	fatturato e ordinativi dell'industria - totale (ATECO CB)		
13	fatturato e ordinativi dell'industria - componente estera - ATECO CI		
14	vendite al dettaglio nella forma distributiva della grande distribuzione specializzata		

Fonte: elaborazione propria

In sintesi, l'approccio *ad-hoc* ha selezionato 14 indicatori da 7 differenti domini: produzione industriale (1), commercio estero (4), margini commerciali (1), vendite al dettaglio (1) fatturato dei servizi (1), fatturato e ordinativi dell'industria (4) ed infine sulle spese per consumi delle famiglie (2).

La regressione implementata con il metodo lasso invece ha selezionato 10 indicatori da 6 domini: produzione industriale (1), commercio estero (6), margini commerciali (1), fatturato e ordinativi dell'industria (1) ed infine dall'indagine sulle spese per consumi delle famiglie (1).

Pur essendo essenzialmente i medesimi domini, i due metodi hanno selezionato in maniera differenziata gli indicatori: l'approccio *ad-hoc* basato sulla correlazione ha optato per quelli relativi all'industria ed ai servizi mentre il lasso in aggiunta a tali serie, ha selezionato anche gli indicatori connessi ai servizi ricreativi.

Gli indicatori così selezionati sono stati sintetizzati mediante due differenti approcci (fase 3): una regressione per componenti principali (ACP) che è stata utilizzata per combinare le serie selezionate con il metodo *ad-hoc*; ed una analisi delle serie individuate con il metodo lasso basata sulla rappresentazione tramite variabili canoniche (ACC).

Le tecniche implementate individuano nuove variabili o fattori che possono essere utilizzati come regressori nelle equazioni per le previsioni oppure – come nel presente lavoro – nel modello ADL(1,1) per la disaggregazione temporale.

In sintesi, con la ACP applicata ai 14 indicatori selezionati con l'approccio *ad-hoc* è stato possibile spiegare il 68,6% della varianza totale con due soli fattori (la stessa tecnica applicata ai 10 indicatori selezionati con lasso ha spiegato il 65,8% della varianza totale con i primi quattro fattori). La riduzione ottenuta tramite l'ACP garantisce, infatti, una sintesi che minimizza la perdita di informazione contenuta nelle variabili di partenza misurata in termini di riduzione delle variabilità totale. Con la ACC, invece, è stato possibile studiare le interrelazioni dinamiche tra le serie storiche componenti il set di regressori selezionati con il metodo lasso. In particolare, l'analisi ha condotto all'individuazione di componenti con capacità previsiva superiore al 75%.

Gli indicatori sintetici descritti nella fase precedente, sono stati utilizzati per produrre una stima trimestrale e la tavola 3 riassume le principali diagnostiche calcolate con il modello di disaggregazione temporale. L'esercizio è stato condotto utilizzando cinque differenti proxy:

- 1) indicatore delle vendite al dettaglio;
- 2) le prime 2 componenti principali individuate a partire dai 14 indicatori selezionati con il metodo *ad-hoc*;

- 3) le prime 4 componenti principali individuate a partire dai 10 indicatori selezionati con il metodo lasso;
- 4) la prima componente individuata a partire dagli indicatori selezionati con il metodo lasso e la rappresentazione canonica;
- 5) le 3 variabili più significative individuate con il metodo lasso: una serie relativa alle importazioni di Servizi nel campo dell'arte, dello spettacolo e del tempo libero (extra UE), una serie relativa alle Spese per consumi delle famiglie in Servizi ricreativi ed una relativa ai margini di trasporto.

Tavola 3. Statistiche calcolate nella fase di disaggregazione temporale con il modello ADL(1,1)

	modello 1	modello 2	modello 3	modello 4	modello 5
	<i>indicatore vendite al dettaglio</i>	<i>ad-hoc &amp; ACP</i>	<i>lasso &amp; ACP</i>	<i>lasso &amp; ACC</i>	<i>modello multivariato</i>
# variabili	1	14	10	10	3
# fattori	1	2	4	1	3
Correlazione	0.40	0.85	0.94	0.88	0.91
Errore standard della regressione	293.3	261.9	209.8	242.2	232.4
Log-verosimiglianza	-206.8	-197.6	-191.44	-203.6	-199.2
AIC	11.45	11.29	10.96	11.10	11.12
BIC	11.54	11.49	11.36	11.11	11.41
Durbin-Watson	2.54	1.63	2.27	2.42	2.50
Jarque-Bera	2.11	1.44	1.01	0.14	0.58
Q-Ljung-Box	4.29	1.40	1.26	3.31	8.52
ME var	-0.82	-1.34	-0.95	-1.06	.678
MAE var	1.55	1.41	1.51	1.81	1.91
RMSE	2.19	2.14	2.04	2.30	2.56
AIC, criterio informativo di Akaike		BIC, criterio informativo Bayesiano			
Durbin-Watson, test autocorrelazione primo ordine dei residui		ME var, errore medio di previsione sui tassi di crescita annuali			
Jarque-Bera, test normalità dei residui		MAE var, errore medio assoluto di previsione sui tassi di crescita annuali			
Q-Ljung-Box, test autocorrelazione di ordine p dei residui		RMSE, errore medio quadratico di previsione sui tassi di crescita annuali			

Fonte: elaborazione propria

L'accuratezza della disaggregazione temporale ottenuta con i diversi modelli e la rispettiva performance in fase previsiva, è stata valutata in funzione della distanza media in valore assoluto tra il tasso di crescita della variabile dipendente (serie annuale di contabilità nazionale) e quello della serie trimestrale annualizzata utilizzando il MAE sui tassi di crescita come misura di riferimento (fase 4).

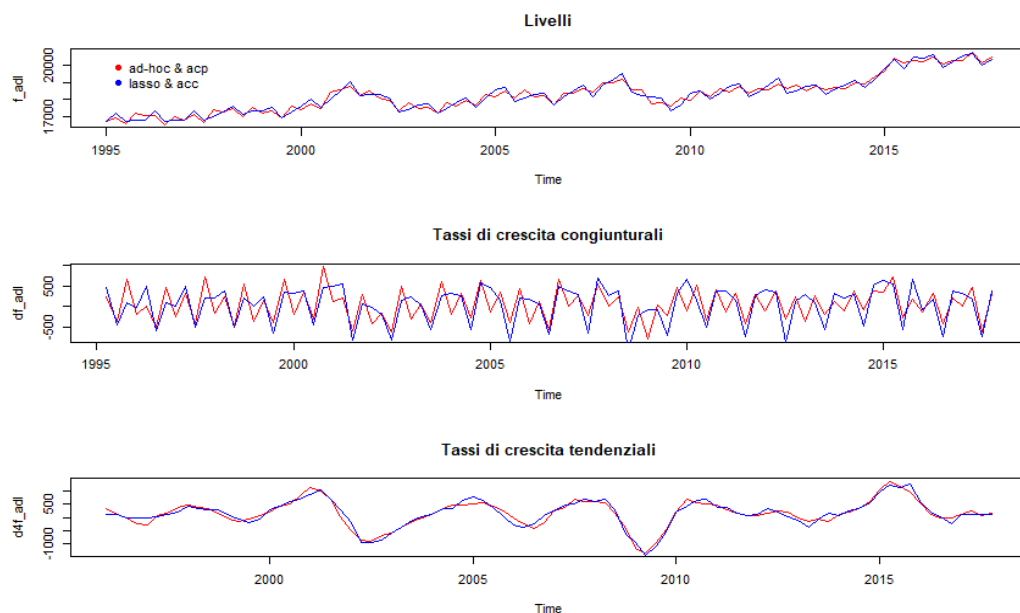
I due approcci rappresentati nei modelli 2 e 4 conducono essenzialmente a dinamiche simili della tendenza di fondo tuttavia il MAE calcolato sul periodo 2010-2017 con l'approccio *ad-hoc/ACP* è leggermente più contenuto e pari a 1,41; mentre l'approccio ottenuto con il metodo *lasso/ACC* è pari a 1,81; pertanto l'indicatore sintetico costruito con l'approccio *ad-hoc/ACP* è – al momento – utilizzato come input delle procedure di trimestralizzazione e destagionalizzazione correntemente adottate nei CET sia per il rilascio a t+30 e t+60 giorni dal trimestre di riferimento per la stima del valore aggiunto del commercio al dettaglio a partire dalla recente revisione dei Conti economici.

La figura 1 mostra tre grafici costruiti sulle stime trimestrali ottenute con un modello di tipo  $ADL(1,1)$  con i differenti approcci esaminati. La principale differenza che emerge analizzando nello specifico il grafico dei tassi di crescita congiunturali, è la differente componente stagionale presente nelle serie trimestrali grezze stimate a partire dai due approcci pur partendo da serie con andamento tendenziale simile. Tali risultati derivano dal fatto che i due approcci partono da indicatori diversi con impatto sull'andamento delle serie.

L'esercizio ha evidenziato quindi come talune scelte nel processo di elaborazioni dei dati ad alta frequenza possano condizionare l'identificazione della componente stagionale, si veda Moauro e Proietti (2008). Per completezza di analisi, il modello 1 costruito sull'indicatore relativo alle vendite al dettaglio è stato scartato a priori in quanto la variabile non è risultata significativa in termini di correlazione (fase 1). Il tentativo di costruire modelli multivariati che tenessero conto di tutti i regressori individuati con i due metodi *lasso/approccio ad-hoc* (rispettivamente 10 e 14) non ha condotto a risultati soddisfacenti, producendo stime trimestrali con elevata volatilità tuttavia limitandosi alle prime tre variabili selezionate con *lasso* è stato costruito un modello con performance soddisfacente (modello 5).

Il modello costruito con le variabili individuate con *lasso* e perequate con l'*ACP*, ha condotto sostanzialmente in termini di diagnostica ai medesimi risultati dell'approccio definito *ad-hoc*.

Figura 1. Serie trimestrali ottenute con i diversi approcci (ad-hoc/ACP, lasso/ACC)  $f\_adl$ , serie nei livelli;  $df\_adl$ , serie dei tassi di crescita congiunturali;  $d4f\_adl$ , serie dei tassi di tendenziali



Fonte: elaborazione propria

## 6. Conclusioni

I risultati raggiunti hanno consentito in prima battuta un approfondimento nella conoscenza del settore economico attraverso l'implementazione degli aspetti metodologici peculiari per grandi mole di dati. Il principale risultato ottenuto è la disponibilità di una stima trimestrale del valore aggiunto dei servizi più dettagliata – settore al momento meno coperto dal regolamento europeo sulle statistiche congiunturali. L'impostazione dell'esercizio può tuttavia contribuire al miglioramento anche di altri settori di attività economica in termini di performance previsiva. Oltre ad estendere l'analisi ad altri settori, l'esercizio potrà essere migliorato quando saranno disponibili ulteriori dati annuali ed introducendo – con serie più lunghe – procedure di validazione più rigorose.

Le tecniche di *variable selection* qui implementate costituiscono, infatti, uno strumento utile alla valutazione di indicatori alternativi favorendo al massimo lo sfruttamento dell'informazione congiunturale disponibile in presenza di carenza o di ritardo nella fornitura degli indicatori ufficiali.

Il lavoro ha proposto di verificare ipotesi diverse basate su differenti concetti: la capacità predittiva e la correlazione. Il modello costruito con due variabili ottenute con

approccio definito *ad-hoc/ACP* è stato ritenuto ottimale rispetto al *lasso/ACC* ed è stato implementato nella procedura di stima corrente dei Conti economici trimestrali.

Nello specifico questo lavoro ha contribuito durante i lavori di revisione dei Conti economici al contenimento degli errori di previsione della stima trimestrale del valore aggiunto del commercio al dettaglio e di conseguenza dei servizi che confluiscono nel Pil trimestrale.

Il lavoro ha inoltre messo in evidenza la difficoltà ad inferire la componente stagionale di una serie partendo dalla frequenza annuale: presi due indicatori, annualizzati e fortemente correlati con la serie annuale di contabilità nazionale, questi potrebbero teoricamente condurre alla rappresentazione del fenomeno in maniera differente, pur conducendo ad un medesimo risultato in termini di capacità predittiva e correlazione. Pur restando un problema di difficile identificazione, il programma di trasmissione di Eurostat relativo alla compilazione dei Conti economici trimestrali richiede ai paesi membri la produzione di serie “grezze” con tutte le incertezze che ne conseguono.

### Ringraziamenti

Sono grata innanzitutto al Professor F. Battaglia per le preziose indicazioni, al collega L. Fenga per il supporto attivo durante l’implementazione dell’esercizio ed al coordinatore dei Conti economici trimestrali italiani F. Moauro con il quale ho condiviso banca dati e scelte metodologiche.

### Bibliografia

- AKAIKE, H. (1974). *Markovian Representation of Stochastic Processes and its application to the analysis of autoregressive moving average processes*, Springer.
- BATTAGLIA, F. (2007). *Metodi di previsione statistica*, Springer.
- BATTAGLIA, F. and FENGA, L. (2003). *Forecasting composite indicators with anticipated information: an application to the industrial production index*, Applied Statistics.
- BISIO, L. and MOAURO, F. (2018). *Temporal disaggregation by dynamic regressions: Recent developments in Italian quarterly national accounts*, Statistica Neerlandica.
- BOX, G. E. P. and TIAO, G. C. (1977). *A canonical analysis of multiple time series*, Biometrika.
- BRILLINGER, D. R. (1981). *Time Series Data Analysis and Theory*, expanded edn. San Francisco.
- CHOW, G. C. and LIN, A.I. (1971). *Best linear unbiased interpolation, distribution, and extrapolation of time series by related series*. The review of Economics and Statistics (pp. 372-375).
- EUROSTAT (2018). *Analysis of the most recent modelling techniques for big data*.
- IMF (2017). *Big Data: Potential, Challenges, and Statistical Implication*.

- FERNANDEZ, R. B. (1981). *A methodological note on the estimation of time series*. The Review of Economics and Statistics, No. 63 (pp. 471-476).
- HOTELLING, H. (1933). *Analysis of a complex of statistical variables into principal components*. Journal of Educational Psychology, No. 24 (pp. 417-441).
- ISTAT (2015). *Conti economici trimestrali. Principali elementi informativi*. [nota aggiornata (2024)].
- MOAURO, F. e PROIETTI, T. (2008). *Temporal Disaggregation and the Adjustment of Quarterly National Accounts for Seasonal and Calendar Effects*, Journal of Official Statistics.
- PEARSON, K. (1901). *On lines and planes of closest fit to systems of points in space*. Philosophical Magazine 2, No. 11, (pp. 559-572).
- PROIETTI (2006). *Temporal disaggregation by state space methods: Dynamic regression methods revisited*. The Econometrics Journal, No. 9 (pp. 357-372).
- STOCK e WATSON (2002). *Forecasting Using Principal Components from a Large Number of Predictors*. Journal of the American Statistical Association Vol. 97, No. 460.
- TIBSHIRANI, R. (1996). *Regression Shrinkage and Selection via the Lasso*, Journal of the Royal Statistical Society. Series B Vol. 58, No. 1 (pp. 267-288),



## INDICATORI DI ALLERTA PER LA PREVENZIONE DELLE CRISI D'IMPRESA: UN'ANALISI STATISTICA SULLE PMI LOMBARDE

### **WARNING INDICATORS FOR BUSINESS CRISIS PREVENTION: A STATISTICAL ANALYSIS OF LOMBARDY SMEs**

*Eugenio Brentari<sup>1</sup>, Gloria Ermici<sup>2</sup>, Andrea Zanetti<sup>3</sup>*

#### **Sommario**

Nel 2019 il Legislatore italiano ha introdotto una nuova versione del Codice della Crisi e dell'Insolvenza, che include una serie di "indicatori di allerta" in grado di individuare e prevenire tempestivamente lo stato di crisi delle imprese. Le imprese che presentano tutti gli indicatori oltre il limite sono tenute a comunicare lo stato di crisi inizialmente alla Camera di Commercio e ora presso il Tribunale competente.

Gli indicatori previsti dalla legge italiana sono strumenti adeguati per prevedere la crisi delle aziende?

Questa ricerca sviluppa un dataset includendo circa 26.000 PMI con sede in Lombardia ed operanti nel settore manifatturiero con l'obiettivo di verificare se tali indicatori sono strumenti idonei per la previsione della crisi d'impresa, anche confrontandoli con il modello Z-score di Altman, uno dei principali modelli predittivi dello stato di crisi aziendale.

Nel nostro studio, questi indicatori non sembrano essere gli strumenti migliori per prevenire lo stato di crisi aziendale.

#### **Abstract**

*In 2019, the Italian legislature introduced a new version of the Crisis and Insolvency Code, which includes a series of "warning indicators" capable of promptly identifying and preventing business crisis. Companies with all indicators exceeding the threshold are required to report the crisis, initially to the Chamber of Commerce and now to the competent Court of Law.*

*Are the indicators required by Italian law adequate tools for predicting business crisis?*

<sup>1</sup> Università degli Studi di Brescia, Dipartimento di Economia e Management, Brescia, Italia - e-mail: eugenio.brentari@unibs.it

<sup>2</sup> Università degli Studi di Brescia, Dipartimento di Economia e Management, Brescia, Italia - e-mail: gloria.ermici@unibs.it

<sup>3</sup> Università degli Studi di Brescia, Dipartimento di Economia e Management, Brescia, Italia - e-mail: andrea.zanetti@unibs.it

*This research develops a dataset including approximately 26,000 SMEs based in Lombardy and operating in the manufacturing sector. The aim is to verify whether these indicators are suitable tools for predicting business crisis, also comparing them with the Altman Z-score model, one of the main predictive models of business crisis.*

*In our study, these indicators do not appear to be the best tools for preventing business crisis.*

## 1. Introduzione

La legge fallimentare italiana ha da sempre fornito strumenti per intervenire sulle imprese già in stato di crisi anziché creare un meccanismo di allerta che fosse in grado di prevedere l'insorgenza di tale fenomeno.

Tale concetto era stato già affrontato nel dettaglio da vari autori con l'obiettivo di creare valutazioni o strumenti predittivi che fossero appunto in grado di analizzare la continuità aziendale e di anticipare la crisi aziendale (Chye Koh e Kee Low, 2004).

Il "principio della continuità aziendale" è alla base del bilancio. Secondo il principio di revisione internazionale ISA 570, esso prevede che il bilancio sia redatto nel presupposto che l'impresa sia in funzionamento e che continuerà la sua attività nel prevedibile futuro.

Una delle conseguenze della mancanza di attenzione al concetto della continuità aziendale è lo squilibrio economico e finanziario dell'azienda, aspetto che porta la società al fallimento. Per esempio, secondo i dati ISTAT, il numero complessivo di fallimenti in Italia nel terzo trimestre del 2023 è aumentato del 5,4% rispetto al trimestre precedente e dell'11,4% rispetto al terzo trimestre del 2022.

Il modello Z-score di Altman è uno dei modelli predittivi più noti a livello internazionale (vedi Altman, 1968, 1970, 1993 e 2002; Altman e Hotchkiss, 1995) e molti autori lo hanno successivamente adattato alla prospettiva italiana caratterizzata da piccole e medie imprese (Alberici, 1975; Altman, Falini e Danovi, 2013).

Il Decreto Legislativo 14/2019 relativo alla riforma della legge fallimentare e al Codice della Crisi e dell'insolvenza ha introdotto il concetto della rilevazione previsionale dello stato di crisi delegando al Consiglio Nazionale dei Dottori Commercialisti ed Esperti Contabili (CNDCEC) il compito di elaborare indicatori di allerta in grado di individuare e prevenire tempestivamente lo stato di insolvenza delle imprese nel rispetto della continuità aziendale.

Il CNDCEC ha determinato cinque indicatori definendo che, nel caso in cui tutti e cinque gli indicatori non risultassero nei limiti previsti dalla normativa, l'azienda ha l'obbligo di comunicare lo stato di crisi inizialmente alla Camera di Commercio e successivamente al Tribunale.

L'applicazione del Nuovo Codice della Crisi, introdotto nel 2019 e successivamente

modificato, ha ampliato la responsabilità dell'amministratore, comportando la necessità di adottare adeguati assetti organizzativi, amministrativi e contabili idonei a rilevare situazioni di insolvenza e ad adottare elementi correttivi.

A luglio 2022 il Nuovo Codice della Crisi è stato modificato, introducendo diversi indicatori, focalizzati su cassa e debiti. In ogni caso, gli indicatori precedenti rimangono uno strumento a disposizione delle aziende per verificare il proprio assetto organizzativo, amministrativo e contabile.

La presente ricerca sviluppa un *dataset* con l'obiettivo di analizzare il comportamento delle aziende e di verificare se gli indicatori previsti dalla legge italiana siano strumenti idonei a prevedere il fallimento delle aziende.

La ricerca analizza successivamente il modello *Z' Score* di Altman sul campione determinato, con l'obiettivo di verificare se esiste una correlazione tra i risultati ottenuti dai due modelli e quale modello sia preferibile per prevedere l'insolvenza delle imprese.

## 2. Indicatori

La combinazione degli indicatori è stata determinata dal CNDCEC<sup>4</sup> con la collaborazione della Direzione Centrale Bilanci Cerved<sup>5</sup> con l'obiettivo di identificare degli indici che fossero rappresentativi di squilibri reddituali, patrimoniali o finanziari nelle aziende.

Per ogni indicatore sono state definite soglie critiche differenti a seconda dei diversi settori di appartenenza delle aziende.

Di seguito sono riportati gli indicatori di allerta predittiva con le soglie identificate per le aziende che operano nel settore manifatturiero:

Oneri finanziari/ Ricavi delle Vendite e delle prestazioni < 3,0%

Patrimonio Netto / Totale Attivo > 7,6%

(Attivo Corrente + Liquidità e mezzi equivalenti) / Debiti a breve termine > 93,7%

Cash flow / Totale Attivo > 0,5%

Debiti fiscali e previdenziali / Totale Attivo < 4,9%

## 3. Modello *Z' Score* di Altman

Il modello *Z' Score* di Altman è uno dei modelli predittivi più conosciuti a livello internazionale. Il modello, basato su una relazione lineare, considera due gruppi societari di pari ampiezza e identifica cinque variabili che devono essere misurate per ciascun

<sup>4</sup> Consiglio Nazionale dei Dottori Commercialisti ed Esperti Contabili, *Crisi d'impresa – Gli indici dell'al-lerta*, 20/10/2019. [http://press-magazine.it/wp-content/uploads/2019/10/codice-della-crisi\\_definizioni-indici\\_ott-2019.pdf](http://press-magazine.it/wp-content/uploads/2019/10/codice-della-crisi_definizioni-indici_ott-2019.pdf).

<sup>5</sup> Cerved Group S.p.a. è una delle aziende più note in Italia per i dataset aziendali <https://www.cerved.com/>.

elemento del campione e ponderate tramite dei coefficienti. Nello studio, Altman identifica cinque indici di bilancio rappresentativi della liquidità, della redditività, del debito e della solvibilità delle aziende.

Successivamente il modello iniziale è stato adattato alle piccole e medie imprese ed è stato sviluppato il modello *Z' Score* (Altman, 1993) che si configura, in termini analitici, con la seguente funzione lineare:

$$Z' = 0,717 X_1 + 0,847 X_2 + 3,107 X_3 + 0,420 X_4 + 0,998 X_5$$

Dove gli indici indicano rispettivamente:

- |  |  |
|--|--|
| $X_1$ : Capitale circolante / Capitale investito | $X_2$ : Utili non distribuiti / Capitale investito |
| $X_3$ : EBIT / Capitale investito                | $X_4$ : Patrimonio Netto / Debiti totali           |
| $X_5$ : Vendite / Capitale investito             |  |

Il modello *Z' Score* di Altman identifica le seguenti tre aree:

<i>Zona di insolvenza</i>	<i>Zona di incertezza</i>	<i>Zona senza rischio</i>
$Z' < 1,23$	$1,23 < Z' < 2,90$	$Z' > 2,90$

#### 4. Il dataset

Il *dataset* prelevato da *Analisi Informatizzata delle Aziende Italiane (AIDA)* include i dati relativi a circa 26.000 piccole e medie imprese (PMI) con sede in Lombardia.

Per la selezione delle PMI sono stati considerati i limiti contenuti nella definizione riportata nella Raccomandazione della Commissione Europea del 6 maggio 2003 introducendo ulteriori requisiti come la sede legale in Lombardia e la disponibilità di bilanci relativi al 2023.

Inoltre, le aziende comprese nel campione fanno riferimento al gruppo merceologico – Manifattura relativo ai codici ATECO appartenenti al gruppo C.

Di seguito si riportano i limiti dimensionali che sono stati considerati per la selezione del dataset:

- Ricavi delle Vendite e delle prestazioni (A) < 50.000.000 Euro
- Dipendenti < 250
- Totale Attivo < 43.000.000 Euro

Ogni record considera i dati economici e finanziari del bilancio 2023 pubblicato dalle Camere di Commercio della Lombardia. In primo luogo, è stato verificato per le società del campione che il patrimonio netto fosse > 0, un indicatore fondamentale per determinare la presenza o meno della continuità aziendale all'interno delle Società.

Partendo dallo stesso *dataset* selezionato per gli indicatori di allerta è stato calcolato

l'indice *Z' Score* di Altman classificando le aziende nelle tre diverse aree precedentemente specificate (zona di insolvenza, zona di incertezza e zona senza rischio).

Come prima analisi è stato confrontato quanto emerso dall'analisi degli indicatori e dall'indice *Z' Score* con l'elenco:

- delle aziende fallite, presso i Tribunali della Lombardia<sup>6</sup>;
- delle aziende inattive o cessate, presso le Camere di Commercio della Lombardia.

Successivamente sono stati confrontati i risultati emersi dall'analisi degli indicatori con quelli ottenuti con l'indice *Z' Score*.

I dati analizzati per il calcolo degli indicatori sono riferiti all'anno 2023, i dati relativi ai fallimenti sono riferiti alle procedure concorsuali avviate a partire dal 01/01/2024 fino al 31/03/2025.

## 5. Risultati

La Tabella 1 mostra i risultati del primo indicatore di allerta (Patrimonio netto > 0) e degli ulteriori 5 indicatori.

*Tabella 1. Risultato dall'analisi degli indicatori*

<b>Indicatori</b>	<i>Totale Aziende nel campione (con i dati disponibili) (A)</i>	<i>Numero di Aziende fuori soglia (B)</i>	<i>% di incidenza (B/A)</i>
Patrimonio netto >0	26.346	1.911	7,25
Oneri finanziari/ Ricavi delle Vendite e delle prestazioni < 3.0%	24.839	2.592	10,44
Patrimonio Netto / Totale Attivo > 7.6%	26.268	3.649	13,89
(Attivo Corrente + Liquidità e mezzi equivalenti) / Debiti a breve termine > 93.70%	26.221	3.565	13,53
Cash flow / Totale Attivo > 0.5%	26.338	3.958	15,02
Debiti fiscali e previdenziali / Totale Attivo < 4.9%	4.038	514	12,73

I dati mostrano che per 1.911 aziende il primo elemento pregiudizievole della continuità aziendale non è rispettato. Secondo la legge italiana queste aziende (corrispondenti a circa il 7,25% del campione totale) dovrebbero avviare la procedura di allerta presso il Tribunale.

Le aziende residuali hanno l'obbligo di verificare gli indicatori successivi e quanti di questi sono fuori soglia, come riportato nella Tabella 2.

<sup>6</sup> I dati includono i fallimenti dall'01/01/2024 al 31/03/2025 dichiarati dai Tribunali della Lombardia.

Tabella 2. Indicatori fuori soglia

Numero degli indicatori fuori soglia	Aziende (A)	% di incidenza (A/24.839)
0	17.545	71,80
1	4.864	18,46
2	1.823	6,92
3	999	3,79
4	902	342
5	204	0,77

Le 204 Aziende con tutti e 5 gli indicatori fuori soglia hanno l'obbligo di avviare la procedura di allerta inizialmente presso la Camera di Commercio e successivamente presso il Tribunale. Queste aziende si sommano a quelle già individuate nella Tabella 1 con il patrimonio netto minore di zero e pertanto, il totale delle aziende che dovrebbero avviare la procedura presso il Tribunale è pari a  $204 + 1.911 = 2.115$  aziende. Secondo l'applicazione di tale disposizione, lo 0,77% del campione totale analizzato dovrebbe quindi avviare la procedura di allerta.

La Tabella 3 mostra i risultati dell'analisi dell'indice *Z' Score*:

Tabella 3. Risultati dell'indice *Z' Score*

Indicatori	Numero di Aziende (A)	% di incidenza (A/24.839)
Zona di insolvenza ( $Z' < 1,23$ )	5.302	20,18
Zona di incertezza ( $1,23 < Z' < 2,90$ )	14.123	53,77
Zona senza rischio ( $Z' > 2,90$ )	6.843	26,05

Con l'applicazione dell'indice *Z' Score*, il 20,18% del campione presenta un rischio di insolvenza.

Sorge allora spontanea la domanda: quante di queste 5.302 aziende hanno effettivamente cessato l'attività o sono fallite?

Considerando l'elenco delle *società fallite* dei Tribunali della Lombardia le aziende dichiarate fallite dai Tribunali della Lombardia nel periodo compreso tra il 01/01/2024 e il 31/03/2025 sono state 1.772, di cui solo una parte era presente nel campione oggetto di analisi.

Considerando invece l'elenco delle *aziende inattive o cessate* delle Camere di Commercio della Lombardia, si sono potute analizzare le aziende inattive o cessate nel pe-

riodo 01/01/2024-31/03/2025 e i loro indicatori fuori *range* con le relative risultanze dell'indice *Z' Score*.

Le aziende inattive o cessate in quel periodo e incluse nel campione esaminato erano 1.431, così suddivise per numero di indicatori fuori *range* (Tabella 4).

Tabella 4. Aziende inattive o cessate con indicatori fuori range

<b>Numero di indicatori fuori intervallo</b>	<b>Aziende (A)</b>	<b>% del campione (A/1.431)</b>
0	236	16,49
1	363	25,37
2	145	10,13
3	209	14,61
4	407	28,44
5	70	4,89
5 + patrimonio netto <0	1	0,07

Le aziende inattive o cessate in quel periodo secondo l'indice *Z' Score* di Altman sono riportate invece nella Tabella 5. Il campione varia da 1.431 a 1.393 unità in quanto i dati di 38 società inattive/cessate non permettono il calcolo dello Z score.

Tabella 5. Aziende inattive o cessate con indice *Z' Score*

<b>Punteggio Z</b>	<b>Aziende (A)</b>	<b>% del campione (A/1.393)</b>
Zona di insolvenza ( $Z' < 1,23$ )	918	65,90
Zona di incertezza ( $1,23 < Z' < 2,90$ )	201	14,43
Zona senza rischio ( $Z' > 2,90$ )	274	19,67

I dati mostrano un numero maggiore di indicatori fuori *range* e di aree di insolvenza nell'elenco relativo a società fallite, inattive o cessate, rispetto al totale delle aziende prese in esame (26.346). I dati sono riportati anche nelle Tabelle 6 e 7.

Tabella 6. Confronto indicatori fuori range campione totale – aziende fallite, inattive o cessate

Numero di indicatori fuori intervallo	% del campione totale (26.346)	% delle aziende fallite, inattive o cessate (1.431)	Differenza
0	66,59	16,49	-50,10
1	18,46	25,37	+6,90
2	6,92	10,13	+3,21
3	3,79	14,61	+10,81
4	3,42	28,44	+25,02
5 + patrimonio netto <0	0,77	4,89	+4,12

Tabella 7. Aziende inattive o cessate con indice Z' Score

Punteggio Z	% del campione totale (26.268)	% delle aziende fallite, inattive o cessate (1.393)	Differenza
Zona di insolvenza ( $Z' < 1,23$ )	20,18	65,90	+45,72
Zona di incertezza ( $1,23 < Z' < 2,90$ )	53,77	14,43	-39,34
Zona senza rischio ( $Z' > 2,90$ )	26,05	19,67	-6,38

Il campione varia da 26.346 (Tabella 6) a 26.268 (Tabella 7) unità in quanto i dati di 78 società non permettono il calcolo dello Z score.

## 6. Conclusioni

Questo articolo analizza i bilanci 2023 delle PMI lombarde, con l'elenco delle aziende fallite dei Tribunali della Lombardia e quello delle aziende inattive o cessate delle Camere di Commercio della Lombardia.

Gli indicatori sono lo strumento giusto per prevedere il fallimento di un'azienda?

Per la disciplina relativa agli indicatori di allerta, se un'azienda ha meno di tre indicatori fuori *range*, può essere considerata “non in stato di crisi” e si può procedere a comparare i risultati ottenuti con i diversi modelli.

I dati mostrano che il modello degli indicatori aiuta a prevedere lo stato di crisi, ma non è sufficiente: nelle 1.431 aziende fallite incluse nel campione, il 51.99% presenta meno di tre indicatori fuori range.

Al contrario, l'indice *Z' Score* è più selettivo e, da questo studio, risulta più appropriato: la zona senza rischio per le aziende fallite rappresenta solo il 19,16% sul totale delle aziende fallite.

Pertanto, si può concludere che il modello *Z' Score* è più selettivo e richiede una maggiore stabilità finanziaria per entrare nella zona a rischio zero, rispetto al modello a indicatori.

Una delle possibili cause potrebbe essere che i modelli a indicatori vengono considerati un obbligo per gli amministratori con responsabilità importanti in caso di inadempienza. Il modello *Z' Score*, invece, è definito come analisi economica e finanziaria.

L'introduzione anche del modello *Z' Score* spingerebbe le aziende a migliorare le proprie competenze di *budgeting* e controllo di gestione, offrendo alle PMI italiane un'opportunità per raggiungere un processo decisionale più efficace. Analizzando i dati, questi modelli possono rappresentare un'importante opportunità per sviluppare analisi predittive nelle PMI italiane, contribuendo alla definizione della strategia e della gestione aziendale.

La ricerca potrebbe essere estesa analizzando altre regioni italiane e con un *dataset* pluriennale oppure introducendo un nuovo e migliore modello indicatore creato dal campione, eventualmente collegato al modello *Z' Score* magari utilizzando tecniche di *machine learning*.

## Bibliografia

- ALBERICI, A. (1975). *Analisi dei bilanci e previsione delle insolvenze: affidabilità bancaria e informativa del mercato mobiliare*, Isedi, Milano
- ALTMAN, E.I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *The Journal of Finance*, Vol. 23, n. 4; pp. 589-609, <https://doi.org/10.2307/2978933>
- ALTMAN, E.I. (1970). Ratio Analysis and the Prediction of Firm Failure: a Reply, *The Journal of Finance*, vol. 25, n. 25, pp. 1169-1172, <https://doi.org/10.2307/2325591>
- ALTMAN, E.I. (1993). *Corporate financial distress and bankruptcy: a complete guide to predicting & avoiding distress and profiting from bankruptcy*, Wiley, New York, ISBN: 978-04-7155-253-6
- ALTMAN, E.I. (2002). Revisiting Credit Scoring Models in a Basel 2 Environment (May 2002). *NYU Working Paper No. S-FI-02-II*, SSRN: <https://ssrn.com/abstract=1298829>
- ALTMAN, E.I., FALINI A., DANOVÌ, A. (2013). Z-Score Models' Application to Italian Companies Subject to Extraordinary Administration (May 1, 2013). *Bancaria* No. 04-2013, SSRN: <https://ssrn.com/abstract=2275390>

- ALTMAN, E.I., HOTCHKISS, E. (1995). *Corporate Financial Distress and Bankruptcy: A Complete Guide to Predicting and Avoiding Distress and Profiting from Bankruptcy*, Wiley, New York, ISBN-13: 978-0-471-69189-1
- BOTTANI, P., CIPRIANI L., SERAO, F. (2004). Il modello di analisi Z-score applicate alle PMI, *Amministrazione & Finanza*, Vol. 1/2004, pp. 50-53
- BRENTARI, E., and ERMICI, G., and ZANETTI, A. (2024). The new preventive alert mechanism in the Italian Crisis Code: A statistical approach for predictive insolvency analysis. Statistics, Technology and Data Science for Economic and Social Development, *Book of short papers of the ASA Bologna Conference – Supplement to Volume 35/3* – Available at doi.org/10.26398/asaproc.0012
- CHYE KOH, H. and KEE LOW, C. (2004). “Going concern prediction using data mining techniques”, *Managerial Auditing Journal*, Vol. 19, No. 3, pp. 462-476
- CONSIGLIO NAZIONALE DEI DOTTORI COMMERCIALISTI ED ESPERTI CONTABILI, *Crisi d'impresa – Gli indici dell'allerta, 20/10/2019*. [http://press-magazine.it/wp-content/uploads/2019/10/codice-della-crisi\\_definizioni-indici\\_ott-2019.pdf](http://press-magazine.it/wp-content/uploads/2019/10/codice-della-crisi_definizioni-indici_ott-2019.pdf) . (Last access: 16/05/2023)
- INTERNATIONAL STANDARD ON AUDITING 570 (Revised) (2016). *Going Concern*. [https://www.ifac.org/\\_flysystem/azure-private/publications/files/ISA-570-\(Revised\).pdf](https://www.ifac.org/_flysystem/azure-private/publications/files/ISA-570-(Revised).pdf) (Last access: 16/05/2023)
- ISTAT (2022), *Annuario statistico italiano 2022*. <https://www.istat.it/storage/ASI/2022/capitoli/C14.pdf>. ISBN 978-88-458-2090-86, Chapter 14 (Last access: 16/05/2023)
- ISTAT (2023), *Registrazione e fallimenti di imprese, I trimestre 2023*. [https://www.istat.it/it/files//2023/05/STATFLASH\\_FALLIMENTI-IMPRESE\\_I-TRIM-2023.pdf](https://www.istat.it/it/files//2023/05/STATFLASH_FALLIMENTI-IMPRESE_I-TRIM-2023.pdf). (Last access: 16/05/2023)
- ITALIAN ACCOUNTING ORGANISM (2018). *OIC 11, Finalità e postulati del bilancio d'esercizio*, marzo 2018. <https://www.fondazioneoic.eu/wp-content/uploads/2011/02/2018-03-OIC-11.pdf> (Last access: 16/05/2023)
- STERLING, R. R. (1968). The Going Concern: An examination, *The Accounting Review*, Vol. 43, No. 3 pp. 481-502

## MULTI-SOURCE ANALYSIS OF LABOR DEMAND DYNAMICS

### ANALISI MULTI-FONTE DELLA DINAMICA DELLA DOMANDA DI LAVORO

*Annalisa Lucarelli<sup>1</sup>, Alessandra Righi<sup>2</sup>, Roberta Rizzi<sup>3</sup>*

#### **Abstract**

*This study explores the potential of integrating multiple data sources to improve the analysis of labor demand trends in Italy. Focusing on the period from 2021 to 2024, we examine and compare three primary sources (administrative and statistical): the daily administrative data from Compulsory Communications (CCs) on actual hirings by Ministry of Labor and Social Policies, the quarterly Job Vacancy Surveys conducted by Istat (JVS/LES), and the monthly Excelsior survey on expected hirings by Unioncamere. By analyzing the unique characteristics of each source and tracking the monthly/quarterly series of hiring and those of quarterly job vacancy rates, we assess the consistency and divergence of labor demand indicators for the total economy and across the two main economic sectors (industry and services). Attention is paid to methodological differences and sector-specific dynamics that may affect the interpretation of results. Our comparative framework enables us to evaluate the convergence of trends and highlight discrepancies, offering valuable insights into the complementarity of these sources. The findings contribute to a more nuanced and comprehensive understanding of labor market signals, supporting evidence-based policymaking in the field of employment.*

#### **Sommario**

Questo studio esplora il potenziale dell'integrazione di diverse fonti di dati per migliorare l'analisi delle tendenze della domanda di lavoro in Italia. Concentrandosi sul periodo 2021–2024, esaminiamo e confrontiamo tre fonti principali (amministrative e statistiche): i dati amministrativi giornalieri delle Comunicazioni Obbligatorie (CO) sulle assunzioni effettive del Ministero del Lavoro e delle politiche sociali, le indagini trimestrali sui posti vacanti condotte da Istat (VELA/GI) e l'indagine mensile Excelsior di Unioncamere sulle assunzioni previste. Analizzando le caratteristiche specifiche di ciascun dataset e osservando le serie mensili/trimestrali dei flussi di assunzioni e quelle trimestrali dei posti vacanti, valutiamo la coerenza e le divergenze degli indicatori della domanda totale di lavoro e per quella nei due principali settori

<sup>1</sup> Istat - Istituto nazionale di statistica, Roma, Italia - e-mail: anlucare@istat.it

<sup>2</sup> Istat - Istituto nazionale di statistica, Roma, Italia - e-mail: righi@istat.it

<sup>3</sup> Istat - Istituto nazionale di statistica, Roma, Italia - e-mail: rizzi@istat.it

economici. Particolare attenzione è rivolta alle differenze metodologiche e alle dinamiche settoriali che possono influenzare l'interpretazione dei risultati. L'approccio comparativo consente di esaminare la convergenza delle tendenze e di evidenziare eventuali discrepanze, offrendo spunti utili sulla complementarità di queste fonti. I risultati contribuiscono a una comprensione più articolata e approfondita dei segnali del mercato del lavoro, a supporto di politiche occupazionali basate su evidenze scientifiche.

**Keywords:** *Sectoral analysis, data integration, labor demand trends.*

## 1. Introduction

Accurately interpreting and forecasting labor demand trends from the employer side is inherently complex when relying on a single data source, due to substantial heterogeneity in data collection methodologies, temporal granularity, sectoral representation, and the underlying conceptual definitions of job openings and hiring intentions. Each source captures a different dimension of labor demand – actual employment flows, unfilled vacancies, or anticipated recruitment – and therefore provides only a partial and potentially biased perspective on broader market dynamics (Lucarelli et al., 2024). In this study, we consider three official sources of labor demand statistics: the Ministry of Labor's administrative Compulsory Communications (CCs)<sup>4</sup>, which record all actual activations, terminations, extensions, and transformations events of employment relationships; the Istat Job Vacancy Survey (JVS/LES), which estimates job vacancy rates by NACE Rev. 2 economic activity sections<sup>5</sup>; and the Excelsior Information System monthly survey, produced by Unioncamere, which collects employers' expectations on future hirings.

We began by examining the characteristics and limitations of these sources and the labor demand indicators that can be derived from them. We then analysed whether the three key statistical datasets provide converging signals on employment trends for the overall economy and its main sectors – despite definitional and methodological disparities – and whether any of them displays earlycycle sensitivity, indicating potential suita-

<sup>4</sup> Compulsory Communication are not open data. Istat receives these micro-data according to a regulation (Legislative Decree 322/1989, 196/2003) that obliges administrative bodies to provide the data, if the source is declared in the National Statistical Plan (PSN).

<sup>5</sup> The job vacancy indicators are obtained by integrating two statistical surveys that form part of a joint data production system. The first is the "Monthly Survey on Employment, Working Hours, Wages and Labour Costs in Large Companies", which is a census-type survey of enterprises with more than 500 employees (LES). The second is the "Quarterly Survey on Vacancies and Hours Worked", which is a sample survey covering enterprises with 1–499 employees (JVS).

bility for nowcasting applications. After exploring multiple approaches to comparison, monthly and quarterly hiring rates were computed as the ratio between the flow of hires and the employment stock from the Istat Labor Force Survey (LFS). To assess the potential of Istat's official quarterly job vacancy rate as a leading indicator, the comparison included both contemporaneous hiring rates and those lagged by one quarter.

The structure of the paper is as follows: Section 2 outlines the characteristics and specificities of the data sources and the methodological framework. Section 3 presents the comparative results, distinguishing between monthly and quarterly series and analysing trends for the overall economy and its two main sectors (Industry and Services). The concluding section summarises the main insights gained and offers directions for future research.

## **2. Data sources and methods**

Labor demand refers to the quantity of workforce that employers are willing to hire during a specific period, based on prevailing wage levels and economic conditions. It is a fundamental aspect of the labor market, shaped by several key drivers, including production performance and trends, demand for goods and services, technological progress, fiscal and monetary policies, and demographic and social dynamics.

A clear understanding of labor demand is essential for evaluating labor market health and anticipating workforce needs across industries and regions. Analysing this variable involves consulting a range of statistical and administrative sources, each offering valuable insights. These enable analysts to monitor both labor market flows and stocks, offering a broad view of labor market dynamics.

In the realm of Italian official statistics on labor market demand, three key sources are noteworthy. The CCs from the Ministry of Labor provide daily data (aggregated for our purposes on a monthly and quarterly basis) on employment flows, including activations, transformations and terminations. The Istat JVS/LES offers quarterly insights into job vacancies, defining a vacancy as an unoccupied position actively being filled by the employer. The Excelsior survey is a continuous monthly study targeting private businesses of all sizes in the industrial and service sectors. A forecasting model uses both sample and administrative data to estimate short-term trends in contract activations.

Despite providing interrelated variables, these sources differ significantly in nature. For this reason, Table 1 summarises their main characteristics, including ownership, source type, subjects required to provide data, coverage, statistical units of analysis, statistical variables and level of detail, frequency of data collection and territorial breakdown, as well as timeliness.

CCs are administrative data, JVS/LES, and Excelsior are surveys. Regarding the entities required to provide data, CCs include all employers within NACE Rev. 2 sections

A-U, JVS/LES survey involves enterprises with at least one employee in the private non-agricultural sector (sections B-S), Excelsior survey covers all companies that were active the previous year and had at least one employee, while excluding public, educational, associative, and agricultural entities. All cover regular and quasi-subordinate employment, excluding self-employment. Another notable difference lies in the unit of analysis: CCs focus on employees and quasi-subordinate jobs, JVS/LES on enterprises, and Excelsior on enterprises and provincial local units.

The level of detail in the collected variables also varies. Istat's JVS/LES survey data are broken down by NACE Rev. 2 economic activity sections at the national level, while CCs data provide detailed information by occupation, economic activity, employment duration for fixed-term jobs, and many other relevant variables referring to the characteristics of the job. Excelsior collects variables on sectoral trends in Italy on total number of expected hires, by month and quarter, type of contract (e.g., fixed-term, permanent), geographical distribution (by province) and also on occupational profiles sought by companies by economic sector with details on the characteristics of expected new hires, including occupational group, business function, specific job roles within each group, expected age of candidates, education level and skills required, hiring difficulties and experience required and preferred qualifications, geographical distribution.

Regarding data collection frequency, timeliness and level of dissemination, CCs provide daily data that are disseminated at regional and national levels, 50 days after the reference quarter; the JVS/LES survey provides quarterly data at the national level, with a flash estimate 45 days and a final estimate 70 days; the Excelsior system collects data continuously and disseminates monthly and quarterly results at t-30 days at the regional and national levels.

In the Excelsior survey, a forecasting model has been developed to enable short-term projections of contract activations, using calibrated estimators based on sample data to improve the forecast of entry flows. These estimates are further refined through the integration of historical series derived from administrative data on firms and employment. The survey is conducted monthly and generates a rolling quarterly forecast: each wave provides estimates for the three months following the survey month, rather than for the entire calendar year as in the past (Unioncamere, 2023).

Table 1. Characteristics of administrative and statistical sources on labor demand

	<b>Compulsory Communications (SISCO system in the Ministry of Labor)</b>	<b>Survey on job vacancies and hours worked (JVS/LES)</b>	<b>Excelsior Information System</b>
<b>Data owner</b>	Ministry of Labor	Italian National Institute of Statistics	Unioncamere
<b>Type of source</b>	<b>Administrative</b> (statistically treated) based on the telematic transmission of communications relating to events of activation, termination, extension, and transformation of employee and quasi-subordinate employment relationships by employers (form UNILAV).	<b>Survey</b> , based on a sample of around 29.000 enterprises. Data are collected separately for enterprises with 1-499 and those with at least 500 employees, respectively, by the Istat JVS survey and the Istat monthly census survey (LES) on large enterprises.	Sample-based <b>survey</b> that gathers forecast data from enterprises on expected hires, combining questionnaire responses with administrative inputs when available.
<b>Reporting entities</b>	Employers: both companies and public institutions (excluding the Armed Forces) and households (for domestic work), residents in Italy. In particular: Private employers (excluding temporary employment agencies in the form UNILAV), domestic employers, and Public Administration.  NACE Rev. 2 sections A-U.	Employers: enterprises with at least one employee in the private non-agricultural sector, resident on Italian territory.  NACE Rev. 2 sections B-S.  Public institutions in public administration, education, and human health (sections O, P, and Q) are not fully covered.	Employers: all companies that were active the previous year and had at least one employee, while excluding public, educational, associative, and agricultural entities.
<b>Coverage</b>	Regular and quasi-subordinate employment are included, while self-employment is excluded.  Temporary work and, partially, the maritime sector are excluded.  Top management figures of public and private companies are excluded.	Jobs for employees (for whom social security contributions are paid) and for apprentices.  Managers are included, with the exception of the top management figures.  Staff with project and collaboration contracts and self-employees are excluded.	Regular and quasi-subordinate employment, excluding self-employment.
<b>Statistical unit of analysis</b>	Employees and quasi-subordinate jobs of Italian/ foreign citizens.  Events of activation, termination, extension and transformation of employee and quasi-subordinate employment relationships.	Enterprises.	Enterprises and provincial local units.
<b>Statistical variables and level of detail</b>	Daily flows of activations and terminations of employment relationships and quarterly flows of jobs.  Average number of activation or termination events per worker.  Workers affected by at least one activation or termination event.  Data are broken down by occupation (permanent/fixed-term), NACE Rev. 2 economic activity section level, and duration of employment (for fixed-term jobs).	Stocks of job vacancies and occupied posts.  Inflow and outflows of occupied posts.  Hours worked.  Data are broken down by NACE Rev. 2 economic activity section level.	The total number of expected hires broken down by month, quarter, contract type, and province. It also covers occupational profiles by economic sector, detailing new hire characteristics such as occupational group, business function, job roles, candidate age, education level, required skills, hiring difficulties, work experience, and preferred qualifications.

<b>Frequency of data collection and territorial detail</b>	Data collected continuously (at each activation, transformation, extension, and termination event)  Disseminated quarterly indicators at the national and regional level.	Quarterly data at the national level.	Data are collected on a continuous basis and released monthly and quarterly.
<b>Timeliness</b>	50 days after the reference quarter.	Estimates 45 (flash estimate) and 70 (final estimate) days after the reference quarter.	Dissemination lag of 30 days, at both the regional (province) and national levels.

In our descriptive analysis, after reviewing various measurement approaches for comparing the three information channels, we constructed monthly and quarterly hiring rates using the series of expected and actual employee hiring flow from Excelsior and CCs, respectively. The denominator for these rates was the stock of employees derived from the continuous LFS (Mlps, Istat, Inps, 2010 and Istat, 2025b). Once converted to a quarterly frequency and considered both contemporaneously and with a one-quarter lag, these hiring rates were compared with the official quarterly Istat vacancy rate (Istat, 2025a).

We then compared the expected monthly hiring rates from Excelsior with the actual monthly hiring rates from CCs, extending the analysis to the quarterly series to assess their relationship with the JVS/LES job vacancy rate. This latter indicator is defined as the number of job vacancies for dependent employment, divided by the sum of vacancies and the stock of dependent workers from the same survey.

The analysis covers the period from January 2021 to December 2024. When comparing the JVS/LES series with Excelsior and CCs data, we applied a one-quarter temporal lag to account for conceptual differences across the datasets: the job vacancy rate captures early-cycle dynamics and employers' hiring intentions, whereas hiring rates reflect expected and realised job commencements. It is important to note that, in the case of Excelsior, the entry rate is based on employers' projections rather than observed hiring data.

These differences suggest that the three sources have varying degrees of sensitivity to the economic cycle, with the job vacancy rate offering a more anticipatory signal compared to realised entries.

### 3. Main results

#### *Monthly comparison between the Excelsior and CCs hiring rate series*

The Compulsory Communications proved to be a highly flexible source for our analysis. The availability of microdata made it possible to obtain data comparable with the other sources mentioned. These presented greater constraints, either due to their design (JVS/LES) or because only aggregated information is accessible (Excelsior)

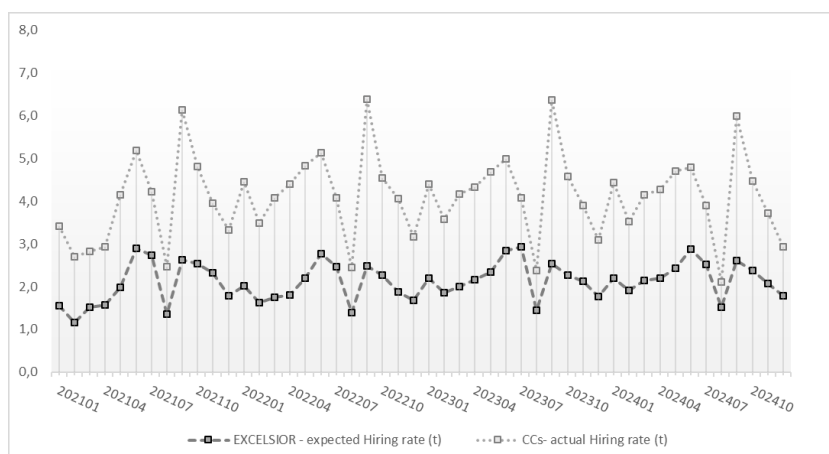
To enhance comparability of trends across the data, particularly at the sectoral level, the enterprise economic activity codes reported in the CCs were integrated with official classifications (based on NACE Rev. 2) drawn from the Istat Business Register (ASIA) for the corresponding reference year. For the 2024 series, the 2023 register was used as it was the most recent available edition.

We constructed monthly hiring rates using the series of expected and actual employee hirings from Excelsior and CCs, respectively. The denominator was the stock of employees from the continuous Labor force survey (LFS)<sup>6</sup>. To ensure consistency, public administration (NACE rev. 2 section O) was excluded from both the CCs numerator and the LFS denominator.

Since the Excelsior survey excludes the entire public sector from its published data on expected hirings, not only Section O, this likely contributes to the lower levels of these hirings compared to actual figures in the total economy and the services (Figures 1 and 3), while similar levels emerge in industry (Figure 2).

The dynamics of expected and actual monthly hiring rates show a high degree of correlation<sup>7</sup>, indicating that peaks and troughs in the economic cycle tend to occur simultaneously in both series. Over the period 2021–2024, the average correlation for the total economy stands at 81.2%, rising 88.6% in the first year of the series (Figure 1), suggesting a strong alignment in hiring patterns.

Figure 1. Expected hiring rate vs. actual hiring rate, Total economy (NACE Rev. 2, sections B-S), 2021M1-2024M12



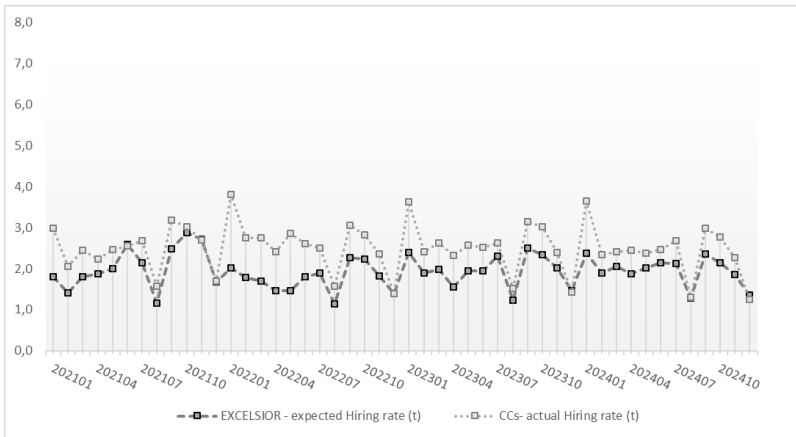
Sources: Excelsior Information System and CCs

<sup>6</sup> For the monthly analysis, these are provisional monthly data.

<sup>7</sup> To analyse the correlation between our indicators, we chose Pearson's correlation coefficient.

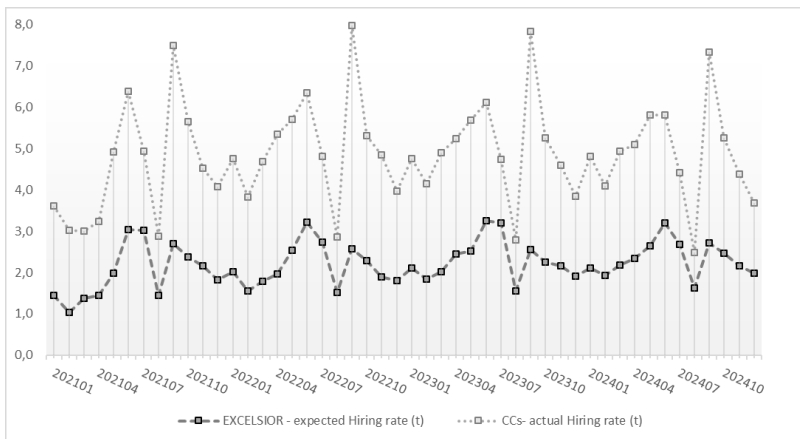
In the sectoral analysis, the average correlation increases in the industrial sector, reaching 83.1% overall and rising to 95.5% in 2024 (Figure 2). By contrast, the service sector exhibits an average correlation of 74.8%, despite reaching 86.2% in 2021 (Figure 3). These results suggest differing degrees of synchrony between expected and actual hiring patterns across sectors. Such differences may stem from the distinct sectoral coverage of the two sources, or from the greater exposure of the service sector to short-term and unforeseen fluctuations.

Figure 2. Expected hiring rate vs. actual hiring rate, Industry (NACE Rev. 2, sections B-F), 2021M1-2024M12



Sources: Excelsior Information System and CCs

Figure 3. Expected hiring rate vs. actual hiring rate, Services (NACE Rev. 2, sections G-S), 2021M1-2024M12



Sources: Excelsior Information System and CCs

### *Quarterly comparison of the Excelsior and CCs hiring rate series with the job vacancy rate*

We now examine the relationship between the job vacancy rate, as reported by JVS/LES, and both expected and actual hiring rates. The job vacancy rate measures the number of vacancies for which search and selection activities are underway, out of 100 vacancies that are either already filled or in the recruitment process. It has been included among the Principal European Economic Indicators (PEEIs), which are used by the European Commission and the European Central Bank to track labor market cycles and broader economic conditions.

In the short term, labor demand tends to follow changes in output. When production rises, firms typically first adjust by extending the working hours of existing staff, both regular and overtime, before proceeding with new recruitments. This creates a time lag between rising output and actual job growth, during which job vacancies emerge. Therefore, vacancies represent an intermediate stage in labor adjustment and act as a leading indicator of cyclical dynamics.

Hiring represents the link between job vacancies and employment, as changes in unmet labor demand affect hiring flows, which in turn affects employment levels.

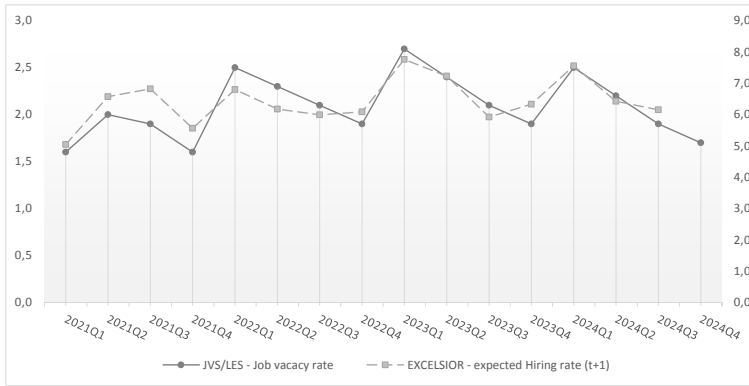
Compared with the extensive literature on the role of job vacancies as a leading indicator of cyclical dynamics (Haggard-Guénette, 1989; Zagorsky, 1998; European Central Bank, 2002; Australian Bureau of Statistics, 2003; Valletta, 2005; Ruth *et al.*, 2006; Mandrone *et al.*, 2010; Lovaglio, 2022), fewer studies have examined the empirical relationship between job vacancies and hiring (Amoah, 2000; Lucarelli, 2010 and 2011).

Building on the evidence that vacancies anticipate employment growth both contemporaneously and one quarter ahead (Lovaglio, 2022), we analysed the average correlations, over the years 2021–2024, between the quarterly vacancy rate series and the expected and actual hiring rates, considering both contemporaneous and lagged by one quarter values.

To ensure consistency in the scope of vacancy and hiring rates, public administration was excluded from both the numerator in the CCs data and the denominator in the LFS data, since JVS/LES and Excelsior both exclude the public sector.

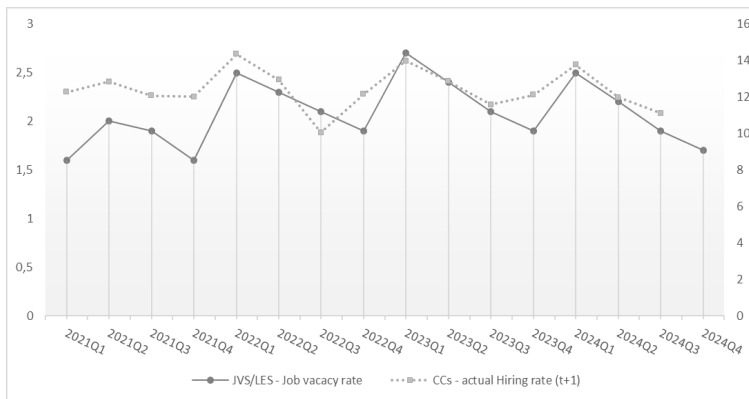
Evidence shows that shifting from contemporaneous to one-quarter lagged relationships resulted in a stronger correlation between job vacancies and hiring rates for the economy as a whole, with values reaching 85% for actual hiring rates and 95% for expected rates. Therefore, as Figures 4 and 5 highlight, the dynamics of the expected and the actual hiring rates (shifted by one quarter) closely correspond to that of the job vacancy rate.

Figure 4. Job vacancy rate vs. expected hiring rate, Total economy (NACE Rev. 2, sections B-S), 2021Q1-2024Q4 (hiring rate one quarter lagged, horizontal axis)



Sources: JVS/LES Istat survey and Excelsior Information System

Figure 5. Job vacancy rate vs. actual hiring rate, Total economy (NACE Rev. 2, sections B-S), 2021Q1-2024Q4 (hiring rate one quarter lagged, horizontal axis)

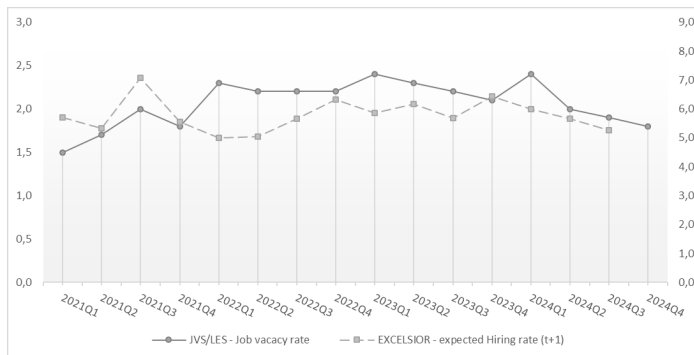


Sources: JVS/LES Istat survey and CCs

The sectoral analysis highlighted different patterns in the industrial and services sectors. In the industrial sector, the average correlation coefficients, computed using hiring rates lagged by one quarter, are lower than those observed for the total economy (reaching around 50%). In some sub-periods, the correlation increases when the simultaneous relationship is considered. One possible explanation is the adoption of more advanced and faster recruitment technologies in industry (e.g., online recruitment channels), particularly among larger firms. Moreover, the provisional nature of monthly hirings of the Excelsior data may translate into concrete search activities (vacancies) within the same quarter.

As shown in Figure 6, the series of Excelsior expected hiring rates lagged by one quarter aligns with the vacancy rates only from the second to the fourth quarter of 2021 and again from the first quarter of 2024 onwards. By contrast a stronger correlation with contemporary Excelsior hiring rates is found in the quarters of 2022 and 2023.

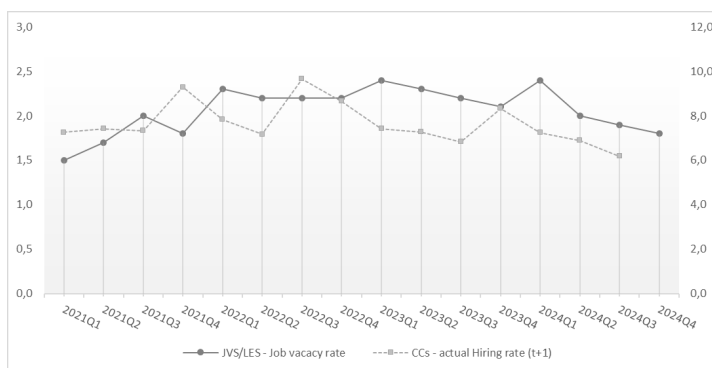
Figure 6. Job vacancy rate vs. expected hiring rate, year-on-year quarterly changes, Industry (NACE Rev. 2, sections B-F), 2021Q1-2024Q4 (hiring rate one quarter lagged, horizontal axis)



Sources: JVS/LES Istat survey and Excelsior Information System

A similar pattern emerges for the CCs actual hiring rate (Figure 7). Signs of alignment between the two dynamics can be observed from the first to the third quarter of 2023, and again throughout the 2024 quarters. As in the previous case, a higher correlation with contemporary variables is observed during the quarter of 2022.

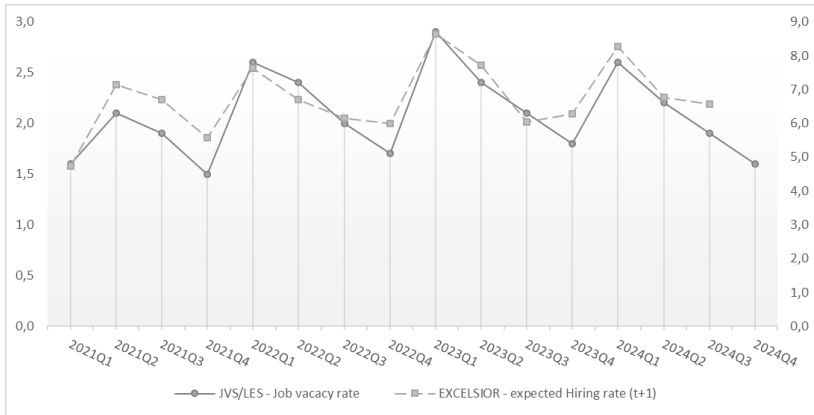
Figure 7. Job vacancy rate vs. actual hiring rate, year-on-year quarterly changes, Industry (NACE Rev. 2, sections B-F), 2021Q1-2024Q4 (hiring rate one quarter lagged, horizontal axis)



Sources: JVS/LES Istat survey and CCs

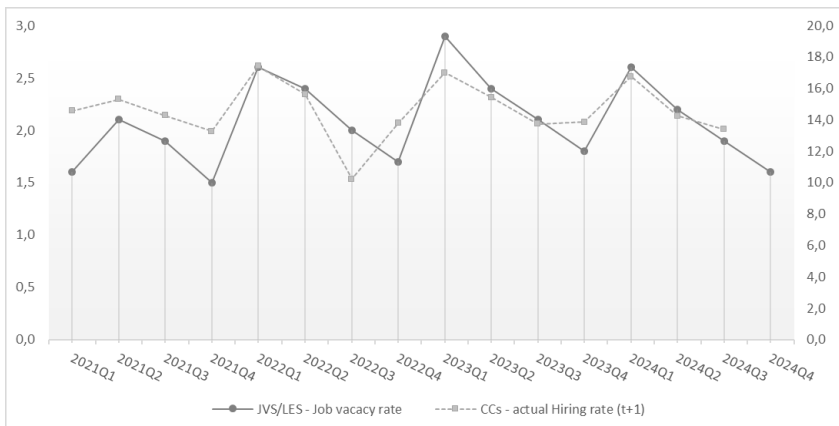
In the services sector (see Figures 8 and 9), both information sources show an even stronger consistency between the vacancy rate series and the lagged hiring rate series than that observed for the economy as a whole (with average annual correlations exceeding those of the latter).

Figure 8. Job vacancy rate vs. expected hiring rate, year-on-year quarterly changes, Services (NACE Rev. 2, sections G-S), 2021Q1-2024Q4 (hiring rate one quarter lagged, horizontal axis)



Sources: JVS/LES Istat survey and Excelsior Information System

Figure 9. Job vacancy rate vs. actual hiring rate, year-on-year quarterly changes, Services (NACE Rev. 2, sections G-S), 2021Q1-2024Q4 (hiring rate one quarter lagged, horizontal axis)



Sources: JVS/LES Istat survey and CCs

#### 4. Conclusions

We compared labor demand dynamics to support a more meaningful interpretation of underlying trends, and the convergence between the rates is encouraging. The combined analysis of Excelsior and Compulsory Communications revealed a notable correlation between expected and actual hiring rates on both a monthly and quarterly basis, particularly in the industrial sector, while the services sector shows a weaker synchronicity. This gap can be attributed to sectoral scope differences between sources, with Excelsior excluding the entire public sector and covering different economic activities compared to CCs. The divergence between forecasts and realised outcomes may also reflect the inherent challenges of forecasting, especially in the services sector, which is more exposed to unpredictable factors such as demand fluctuations, seasonality, and shifts in consumer behavior. Consequently, the capacity to accurately forecast employment flows is more limited than in the industrial sector, where production cycles are typically more structured and stable.

The quarterly comparison with the Job Vacancy Rate confirms its role as a leading indicator of employment, in line with economic literature. The findings show stronger correlations between vacancies and hiring when a one-quarter lag is applied, suggesting that job vacancies precede and influence actual hiring flows. This pattern is particularly evident in the services sector, where lagged dynamics offer useful interpretive insight for reading trends.

In summary, while the results indicate an overall consistency across the data sources, they also reveal distinct sectoral differences, particularly in the alignment and timing of hiring patterns.

Although the indicators are subject to inherent data limitations (such as differences in scope, reporting behaviour, and sectoral coverage), which inevitably influence the precision of crosssource comparisons, the harmonisation efforts undertaken in this study substantially mitigate these constraints and enhance the analytical robustness of the comparison.

Taken together, these findings underscore the importance of integrating multiple data sources and methodological approaches to achieve a more comprehensive and sector-sensitive understanding of labour demand dynamics.

From the perspective of further strengthening this integrative approach – especially with the aim of generating strategically relevant information on labour demand dynamics by occupation and territorial areas – the analytical framework could be extended to include non-traditional data sources, such as online job vacancies (see Lucarelli *et al.*, 2025). This perspective is also consistent with recent proposals advanced by Eurostat advocating the integration of traditional and innovative data sources for the production of experimental labour market indicators (Eurostat, 2024 e 2025). Such efforts highli-

ght the potential of combining administrative data, survey-based statistics, and alternative digital sources to enhance timeliness, granularity, and policy relevance, thereby fostering a more comprehensive and forward-looking understanding of labour market developments.

## References

- AMOAH, B. (2000). Help-wanted index. *Perspectives on Labour and Income*. Statistics Canada, Catalogue no 75-001-XPE, 14-18.
- AUSTRALIAN BUREAU OF STATISTICS (2003). Do Job Vacancies Provide a Leading Indicator of Employment Growth? *Australian Labour Market Statistics*. Catalogue no. 6105.0, April.
- BRANKA, J., KVETAN, V., NAPIERALA, J. (2022). From the online job advertisements to official statistics – The aspects of quality assurance. *Book of Abstracts of the European Conference on Quality in Official Statistics - Q2022*, Vilnius.
- EUROPEAN CENTRAL BANK (2002). *Monthly bulletin*. Frankfurt: European Central Bank.
- EUROSTAT (2024). Which occupations are the most sought on the web? Retrieved from <https://ec.europa.eu/eurostat/en/web/products-eurostat-news/w/ddn-20241210-2>
- EUROSTAT (2025). Which were the most demanded occupations in 2024? Retrieved from <https://ec.europa.eu/eurostat/en/web/products-eurostat-news/w/ddn-20251210-3>
- HAGGAR-GUÉNETTE, C. (1989). Jobs ads: A leading indicator? *Perspectives on Labour and Income*, 1, 2. Statistics Canada, Catalogue no 75-001-XPE, 1-12.
- ISTAT (2025a). Il mercato del lavoro - I Trimestre 2025. Retrieved from <https://www.istat.it/comunicato-stampa/il-mercato-del-lavoro-i-trimestre-2025/>.
- ISTAT (2025b). Occupati e disoccupati (dati provvisori) - Maggio 2025. Retrived from <https://www.istat.it/comunicato-stampa/occupati-e-disoccupati-dati-provvisori-maggio-2025/>.
- LOVAGLIO, P.G. (2022). Do job vacancies variations anticipate employment variations by sector? Some preliminary evidence from Italy, *Labour*, 36(1), 71-93.
- LUCARELLI, A. (2010). Job vacancies and hirings: first evidence coming from enterprise-level analysis. In Italian Statistical Society-SIS (Ed.), *PHD Theses in Statistics and Applications - Book of short papers*, 1, (pp. 159-164). Padova: Cleup.
- LUCARELLI, A. (2011). Vacancies and Hirings: Preliminary Evidence from a Survey on Italian Employers. *Rivista di Statistica Ufficiale*, 2-3, 21-53. Roma: Istat.
- LUCARELLI, A., RIGHI, A. & RIZZI, R. (2025). Improving information on labour demand for decision-making by integrating traditional surveys, administrative sources, and web-based data. In L.S. ALAIMO, F. CRESCENZI, L. FABBRIS, F. MAGGINO (eds.), *Measuring and Interpreting World Changes with Statistics, Data Science and AI*. Book of Short Papers of the ASA Rome

Conference 18 to 20 September 2024 (pp. 293-298). *Italian Journal of Applied Statistics*, Vol. 35/2 Supplement to Vol. 37/2. Firenze: FUP - Firenze University Press. doi.org/10.26398/asaproc.0119

- MANDRONE, E., TANCIONI, M. & LAJ, S. (2010). Vacancies as employment predictor: dynamic properties of the series. In Mandrone, E. (Ed.), *Job vacancies in the Italian labour market. The new ISFOL "Help Wanted" Time Series. Isfol TEMI E RICERCHE*. Soveria Mannelli: Rubettino.
- MLPS, ISTAT & INPS (2010). Nota metodologica per i dati delle Comunicazioni Obbligatorie - *Rapporto di coesione sociale 2010*. Retrived from <https://www.cliclavoro.gov.it/DocumentoMetodologicoSisco.pdf>.
- RUTH, F. J., SCHOUTEN, J. G. & WEKKER, R. F. (2006). Statistics Netherlands business cycle tracer. Methodological aspects; concept, cycle computation and indicator selection. *Statistics Netherlands Discussion Paper*, 5, 1-64.
- VALLETTA, R. (2005). Help-wanted advertising and job vacancies. *Federal Reserve Bank of San Francisco Economic Letter*, 2, 1-3.
- UNIONCAMERE (2023). La metodologia del nuovo sistema informativo Excelsior (anno 2023). Retrived from [https://excelsior.unioncamere.net/sites/default/files/documenti/Nota\\_Metodologica.pdf](https://excelsior.unioncamere.net/sites/default/files/documenti/Nota_Metodologica.pdf).
- ZAGORSKY, J. L. (1998). Job vacancies in the United States: 1923 to 1994. *Review of Economics and Statistics*, 80, 338-345.



# QUANTIFICAZIONE AVANZATA DEL SENTIMENT DEI CLIENTI E PREVISIONE DELLE VENDITE MEDIANTE LARGE LANGUAGE MODELS

## ADVANCING CUSTOMER SENTIMENT QUANTIFICATION AND SALES PREDICTION WITH LARGE LANGUAGE MODELS

*Domenica Fioredistella Iezzi<sup>1</sup>, Roberto Monte<sup>2</sup>, and Daniele Pasquini<sup>1</sup>*

### Sommario

Questo studio esamina l'applicazione dei Large Language Model (LLM) nel campo dell'analisi del sentiment, con particolare attenzione ai contenuti generati dai consumatori relativi alla Toyota Camry nel mercato statunitense. Il corpus è composto da 163.000 tweet in lingua inglese, postati tra giugno 2009 e dicembre 2022. Abbiamo usato tre distinti LLM: CardiffNLP, ChatGPT e LLaMA per la generazione di proxy del sentiment dei consumatori. Ogni modello rappresenta un diverso paradigma metodologico: CardiffNLP è un classificatore supervisionato, specifico per dominio, ottimizzato sui dati di Twitter; ChatGPT è un modello generativo zero-shot di uso generale; e LLaMA è un modello open source basato su prompt, in grado di valutare il sentiment in modo continuo. Forniamo una valutazione comparativa di questi modelli in termini di accuratezza della classificazione, granularità del sentiment e utilità predittiva. I proxy del sentiment vengono quindi incorporati in un modello MARSS (Multivariate AutoRegressive State-Space) per valutarne l'impatto sulla previsione dell'attuale volume di vendite mensili della Toyota Camry. I nostri risultati indicano che il modello MARSS migliora significativamente l'accuratezza delle previsioni rispetto ai tradizionali benchmark ExponentTial Smoothing (ETS), che non possono tenere conto del sentiment, e rispetto a un modello univariato state-space che non considera ancora i segnali del sentiment.

### Abstract

*This study examines the application of Large Language Models (LLMs) in the domain of sentiment analysis, with a specific focus on consumer-generated content related to the Toyota Camry in the U.S. market. By leveraging a corpus of approximately 163,000 English-language tweets collected between June 2009 and December 2022, we examine the performance and consistency of three distinct LLMs: Car-*

<sup>1</sup> Tor Vergata University, Department of Enterprise Engineering "Mario Lucertini", Rome, Italy - e-mail: stella.iezzi@uniroma2.it; daniele.pasquini@uniroma2.it

<sup>2</sup> Tor Vergata University, Department of Civil Engineering and Computer Science Engineering, Rome, Italy - e-mail: roberto.monte@uniroma2.it

*diffNLP, ChatGPT, and LLaMA in generating consumer sentiment proxies. Each model represents a different methodological paradigm: CardiffNLP is a supervised, domain-specific classifier fine-tuned on Twitter data; ChatGPT is a general-purpose, generative zero-shot model; and LLaMA is an open-source, prompt-based model capable of continuous sentiment scoring. We provide a comparative evaluation of these models across classification accuracy, sentiment granularity, and forecasting utility. The sentiment proxies are then incorporated into a Multivariate AutoRegressive State-Space (MARSS) model to assess their impact on predicting the current Toyota Camry's monthly sales volume. Our findings indicate that the MARSS model significantly enhances forecasting accuracy compared to traditional Exponential Smoothing (ETS) benchmarks, which cannot account for sentiment, as well as when compared to a univariate state-space model that does not yet consider the sentiment signals.*

## 1. Introduction

Large Language Models (LLMs) represent a significant advancement in Natural Language Processing (NLP), substantially enhancing machines' ability to comprehend and produce human language. These models are characterised by their large scale, often involving billions of parameters, and by training procedures based on massive and heterogeneous textual corpora (Brown *et al.*, 2020; OpenAI, 2023). At the architectural level, most contemporary LLMs rely on the Transformer framework (Vaswani *et al.*, 2017), which introduced the self-attention mechanism. This mechanism enables models to capture contextual dependencies between words regardless of their distance within a sequence, allowing for a richer representation of syntactic and semantic relationships. As a result, Transformer-based models have become the foundation of state-of-the-art systems such as BERT (Devlin *et al.*, 2019) and successive generations of large generative models. ChatGPT-3 (Brown *et al.*, 2020), and ChatGPT-4 (OpenAI, 2023). LLMs are pre-trained on vast corpora containing diverse text sources, including books, articles, and web pages. Through this process, they acquire broad linguistic knowledge, enabling strong performance across a wide range of downstream tasks, including text summarisation, translation, question answering, and sentiment analysis (Bommasani *et al.*, 2021; Chowdhery *et al.*, 2022). Despite these advances, the application of LLMs in real-world settings presents several challenges. High computational costs affect accessibility and scalability, while issues related to bias, hallucinations, reproducibility, and transparency remain central concerns (Bender *et al.*, 2021). These limitations are particularly relevant in sentiment analysis, where robustness, interpretability, and consistency over time are crucial for downstream analytical tasks such as forecasting and decision support.

In this context, the current study examines the impact of various modelling approaches for sentiment analysis, ranging from domain-specific classifiers to general-purpose and open-source language models, on the extraction and reliability of sentiment signals from social media data.

Our goal is not just to compare predictive accuracy, but also to evaluate the consistency, granularity, and analytical usefulness of sentiment measures derived from different language models when applied to large-scale, real-world datasets. Therefore, we conduct a comparative analysis using three representative sentiment analysis models based on Transformer architectures. These include both large language models (LLMs) and domain-adapted BERT-based models. Our selection reflects different design choices within the same architectural framework rather than fundamentally different methodological approaches. Specifically, the models vary in terms of domain specialization, inference strategy, reproducibility, and accessibility. This variety allows us to systematically evaluate how these design choices influence sentiment quantification and its implications. We apply these models to a dataset of approximately 163,000 tweets related to the Toyota Camry, covering the period from June 2009 to December 2020. From this analysis, we derive alternative sentiment proxies on both discrete and continuous scales, which we subsequently use for sales forecasting.

The structure of this paper is as follows: Section 2 examines LLMs and sentiment quantification; Section 3 describes the corpus and the characteristics of our sentiment; Section 4 presents the main results of both sentiment and forecasting models; Section 5 concludes with implications and directions for future research.

## **2. Large Language Models**

LLMs represent an important leap forward in natural language processing, offering unprecedented capabilities in language understanding and generation. These advances have reshaped the landscape of sentiment analysis, a task traditionally performed using lexicon-based tools and supervised machine learning algorithms that aim to quantify the emotional valence of textual data. Early sentiment analysis approaches, such as VADER (Hutto & Gilbert, 2014; Iezzi & Monte, 2024), relied on predefined sentiment dictionaries and rule-based heuristics. These methods, while effective for short and general-purpose texts, often lacked the nuance needed to handle complex language, sarcasm, or domain-specific terminology. The advent of deep learning and transformer-based architectures marked a shift toward more context-aware methods.

Table 1. Comparison of Prompting and Training Strategies for Language Models

Setting	Examples Provided	Model Trained	Flexibility	Use Case
<i>zero-shot</i>	None	No	High	generalization, fast testing
<i>few-shot</i>	2-5 examples	No	medium-high	more accurate prompt
<i>fine-tuning</i>	Thousands	Yes	low-medium	custom models, high control

Table 1 compares three main approaches for using language models in downstream tasks: zero-shot, few-shot, and fine-tuning. Zero-shot learning requires no examples and no additional training. This method is the most flexible and time-efficient, making it ideal for rapid prototyping, generalization across different domains, and tasks where labeled data is scarce or unavailable. Few-shot learning involves providing a small number of examples directly in the prompt. This approach enhances performance and accuracy over zero-shot learning by offering context while still eliminating the need for model retraining. It effectively balances ease of use and precision. Fine-tuning relies on training the model with large amounts of task-specific data. Although this method provides the highest level of control and customization, it comes with reduced flexibility and increased computational demands. In this study, we selected three representative models for sentiment analysis:

CardiffNLP (Barbieri *et al.*, 2020, 2022a, 2022b), ChatGPT (OpenAI, 2023), and LLaMA (Touvron *et al.*, 2023a, 2023b ; MetaAI, 2023). We based our selection on their complementary characteristics and distinct methodological approaches. Each model has unique advantages in terms of domain specialization, contextual inference, and scalability, enabling a robust comparative evaluation across diverse textual inputs. CardiffNLP is a domain-adapted model built on the RoBERTa architecture and fine-tuned specifically for sentiment classification of tweets. Trained on millions of Twitter posts, it is particularly effective at handling informal, short-form, and noisy text, including emojis, hashtags, and slang. Its performance is optimized for social media analytics, making it a strong baseline for evaluating sentiment in user-generated content. CardiffNLP provides high classification accuracy in the social media domain by leveraging a sentiment polarity scale (positive, neutral, negative).

ChatGPT is a state-of-the-art generative model that excels in zero-shot and few-shot learning scenarios. Unlike domain-specific models, it can infer sentiment across a wide range of contexts, including complex and nuanced language, without needing task-specific fine-tuning. Its ability to interpret sentiment through conversational interaction and capture semantic subtleties makes it ideal for analyzing long-form and ambiguous texts. Additionally, ChatGPT supports flexible outputs, including both discrete and continuous sentiment scales.

LLaMA (Large Language Model Meta AI) serves as an open-source alternative to commercial language models, providing complete transparency and adaptability. Available in several configurations (e.g., LLaMA-2 and LLaMA-3, with up to 70 billion parameters), it is trained on publicly available data rather than proprietary corpora. The LLaMA models are designed for high performance, even in resource-constrained environments, and are suitable for fine-tuning or inference in production workflows. Importantly, LLaMA enables continuous sentiment estimation, extracting sentiment scores on a real-valued scale, which allows for richer downstream analyses, such as forecasting.

*Table 2. Comparison of Sentiment Estimation Methodologies Across Three NLP Models*

Model	Methodology	Training Required	Output Type
ChatGPT-4	Zero-shot (generative)	No	Continuous score
CardiffNLP	Supervised classification	Yes (fine-tuned)	Continuous (via softmax)
LLaMA 3.1-8B	Prompt-based generative	No (unless fine-tuned)	Continuous score

Table 2 compares ChatGPT-4, CardiffNLP, and LLaMA 3.1-8B. It highlights the differences in their methodological approaches, training requirements, and output structures. ChatGPT-4 and LLaMA 3.1-8B are generative, prompt-based models that can produce continuous sentiment scores in zero- or few-shot settings. This property makes them particularly suitable for exploratory analyses and situations with limited annotated data. In contrast, CardiffNLP employs supervised learning and domain-specific fine-tuning, yielding probabilistic sentiment scores across discrete polarity classes. This approach ensures greater stability in structured, annotation-rich environments. These models offer complementary strengths. Generative language models (LLMs) emphasize flexibility and contextual sensitivity, while CardiffNLP provides robust performance in controlled, domain-specific applications.

The sentiment proxies derived by our three language models are then incorporated into a MARSS model (see Holmes, 2024) to assess their impact on predicting the current Toyota Camry's monthly sales volume. In the MARSS model, inspired by the ETS models (see Holt, 1957 and Hyndman *et al.*, 2002) and adopted also as univariate benchmarks when we do not account for consumer sentiment, we assume that the Toyota Camry's monthly sales volume splits into three additive components: a trend component referred to as local level, a seasonal component, and a noise component. In turn, the local level at time  $t$  is updated from the former value at time  $t-1$  by a linear combination of a random walk, referred to as the *slope in the local level*, and an independent innovation. We intend to incorporate consumer sentiment and its observed proxy into the MARSS

model by interpreting the former as the slope in the local level and treating the latter as a second observed state, in addition to the monthly sales volume (see Iezzi & Monte, 2024a, 2024b, 2025).

### 3. Corpus and models

The tweets in the corpus were retrieved using keyword and hashtag-based queries (e.g., “Toyota Camry”, “#Camry”, “@ToyotaCamry”), to ensure diversity in content and temporal coverage (see Iezzi & Monte, 2025). The resulting corpus captures a wide range of expressions, including user experiences, promotional posts, complaints, and humorous remarks, making it particularly well-suited for sentiment analysis in real-world social media contexts. For the annotation with LLMs, we adopted a standardized instruction prompt to ensure comparability across outputs: “*You are a Twitter/X sentiment analysis expert. For each tweet provided, respond with a JSON object containing exactly two keys: “id” and “sentiment\_score”. sentiment\_score must be a float between -1.0 and 1.0 with one decimal place: -1.0 = highly negative, 0.0 = neutral, +1.0 = highly positive. Do not output any additional text*”. To highlight the differences in how each model interprets the emotional content of language, particularly in more nuanced or ambiguous cases, we present an example of the application of the aforementioned sentiment analysis models to the sentences contained in three tweets.

*Table 3. Comparative Sentiment Scoring of the Toyota Camry’s related Tweets using ChatGPT, CardiffNLP, and LLaMA*

Sentence	ChatGPT	CardiffNLP	LLaMA
My new Toyota Camry is smooth, quiet, and super reliable love driving it	0.93	0.99	0.90
The Toyota Camry gets me where I need to go, but it’s not exactly exciting	0.15	-0.49	0.10
The worst decision I made was buying a Toyota Camry constant issues and zero support	-0.89	-0.97	-1.00

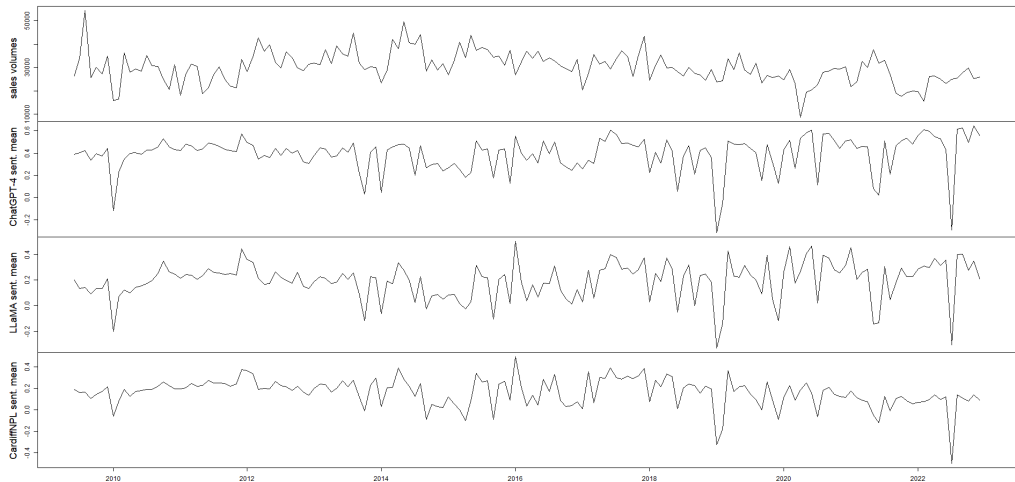
As shown in Table 3, the first sentence, “My new Toyota Camry is smooth, quiet, and super reliable love driving it”, yields consistently high sentiment scores across all three models: ChatGPT assigns a score of +0.93, CardiffNLP +0.99, and LLaMA +0.90. This convergence suggests strong model agreement in detecting the overtly positive tone, driven by explicitly favorable adjectives (“smooth,” “super reliable”) and an enthusiastic emotional appraisal. The second sentence, “The Toyota Camry gets me where I need to go, but it’s not exactly exciting”, reveals some variation in sentiment interpretation. ChatGPT produces a modestly positive score (+0.15), likely influenced by

the phrase emphasizing functionality (“gets me where I need to go”); LLaMA yields more neutral values (+0.10); the CardiffNLP model assigns a moderately negative score (−0.49), suggesting that it gives greater weight to the emotional disappointment than to the functional statement. This case highlights the model’s sensitivity to nuanced negation and contrastive conjunctions in sentiment-bearing clauses. The third sentence, “The worst decision I made was buying a Toyota Camry constant issues and zero support”, results in uniformly negative sentiment predictions. ChatGPT reports −0.89, CardiffNLP −0.97, and LLaMA −1.00. The presence of unambiguous negative language (“worst decision,” “constant issues,” “zero support”) ensures consistent sentiment decoding across all models.

#### 4. Results

The graphs in Figure 1 compare the sentiment scores produced by our LLMs to evaluate both their temporal evolution and the degree of concordance between them and the Toyota Camry’s monthly sales volume.

*Figure 1. Toyota Camry’s Monthly Sales Volume and Consumer Sentiment Signals Proxies across ChatGPT, LLaMA, and CardiffNLP (2009 - 2022)*



More specifically, Figure 1 presents the Toyota Camry’s monthly sales volumes (top panel) alongside the average monthly sentiment scores derived from the applied LLMs. The time frame spans from 2010 to 2023, enabling a long-term comparison between actual consumer behavior (sales) and sentiment signals. The top panel reveals notable seasonality and cyclical patterns in vehicle sales, with a general upward trend until

around 2016, followed by increased volatility and a pronounced dip in 2020, likely attributable to the COVID-19 pandemic. A modest recovery is visible in the post-pandemic period. The second panel shows sentiment estimates generated by ChatGPT. Initially, these values appear to be smoother and less volatile over time compared to the later patterns, indicating a conservative sentiment at the outset. There is some alignment with sales dynamics, especially during the expansion period before 2020 and the decline in 2020. However, in other periods, such as 2012–2013, the correspondence is weaker. The third panel illustrates sentiment signals from LLaMA. This model shows greater responsiveness to changes in textual tone, displaying both stronger positive and negative peaks compared to ChatGPT. However, the sentiment pattern of LLaMA still appears to be closely correlated with ChatGPT. The bottom panel displays sentiment values generated by CardiffNLP. This model, trained as a supervised classifier, tends to produce a slightly smoother signal than ChatGPT and LLaMA, although it impressively exhibits the same pattern as LLaMA. Overall, the figure illustrates how the applied models construct signals with clear commonalities while capturing an underlying message, despite variations in construction techniques. This issue might be due to the model training domain (social media text), which may yield more volatile sentiment estimates over time. The right panel shows the residuals, i.e., the difference between the estimated sentiment signal and its ETS-ANN smoothed trend. ChatGPT residuals exhibit a somewhat more moderate dispersion, with frequent but relatively stable oscillations. Occasional large deviations are observable, particularly during the 2019–2021 period, indicating temporal mismatches between observed sentiment and long-term trends. LLaMA residuals are more dispersed and exhibit higher volatility, consistent with the model's responsiveness to short-term fluctuations and possible overreaction to noise in the underlying text data. CardiffNLP residuals exhibit slower volatility compared to LLaMA, indicating a less noisy signal.

Figure 2. Toyota Camry's Monthly Consumer Sentiment Proxies across ChatGPT, LLaMA, and CardiffNLP (2009 - 2022) - ETS-ANN Decomposition - Fitted Values and Residuals

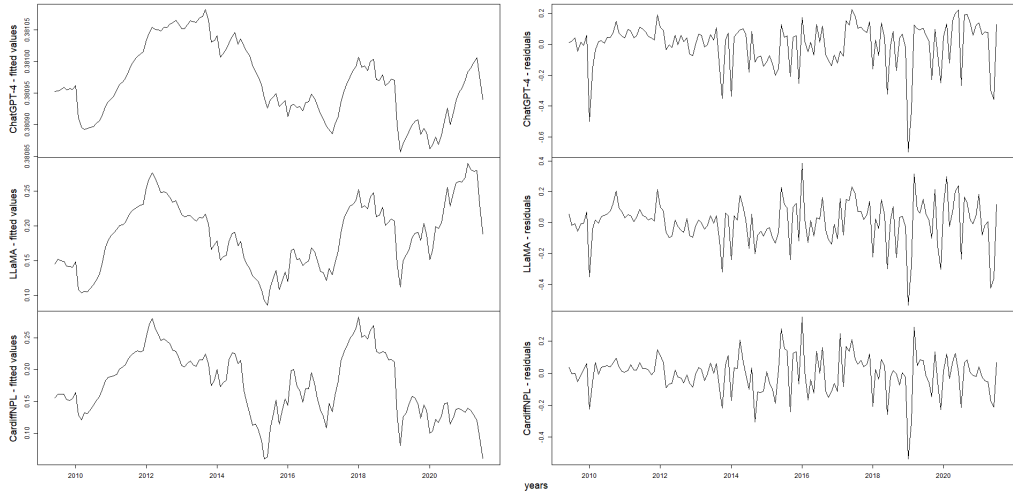
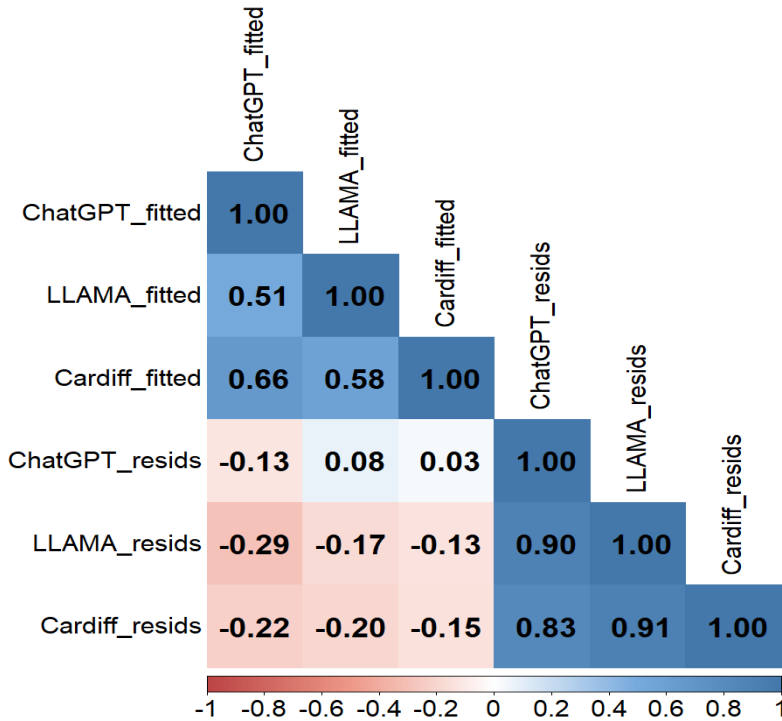


Figure 3 presents the zero-lag correlation matrix for the fitted values and residuals obtained from the ETS-ANN. This matrix provides insight into the coherence of both the long-term trend components and the short-term fluctuations captured by the different sentiment extraction techniques. The fitted values, which represent the smoothed underlying trend of consumer sentiment, exhibit moderate to strong positive correlations across models. ChatGPT and CardiffNLP show the highest correlation ( $\rho = 0.66$ ), followed by CardiffNLP and LLaMA ( $\rho = 0.58$ ), and ChatGPT and LLaMA ( $\rho = 0.51$ ). These results reinforce the graphical evidence that, despite their methodological differences, ranging from generative prompting (ChatGPT, LLaMA) to supervised classification (CardiffNLP), the models capture a shared temporal structure in consumer sentiment over time. The residuals, representing high-frequency variations not explained by the trend, are even more tightly correlated. The correlation coefficients exceed 0.83 for all pairwise comparisons, with the highest observed value between LLaMA and CardiffNLP ( $\rho = 0.91$ ). This strong agreement in the residual component indicates that short-term sentiment fluctuations, likely corresponding to specific events or shocks (e.g., product recalls, marketing campaigns), are detected similarly across models, reinforcing their utility for real-time monitoring applications. Importantly, the cross-correlations between fitted values and residuals are close to zero or mildly negative, as expected under the assumptions of the ETS-ANN framework. This orthogonality validates the

decomposition approach and confirms that the trend and noise components are indeed separated appropriately. These findings underscore the consistency and complementarity of ChatGPT, LLaMA, and CardiffNLP in tracking both structural and transient aspects of consumer sentiment, justifying their joint use in multi-model ensemble strategies for sentiment-based forecasting. Table 3 summarizes some information parameters and accuracy measures for the Toyota Camry's monthly sales volume, both with and without the proxy signal of consumer sentiment. We achieve the most accurate results using the sentiment proxy derived by ChatGPT, LLaMA, and CardiffNLP. In this context, the ETS model with Additive noise, No slope in the local level, and Additive seasonality (ETS-ANN) was still optimally selected from the ETS family to analyze the Toyota Camry's monthly sales volume, without accounting for the consumer sentiment proxy. However, we also wanted to consider the slightly suboptimal ETS model with Additive noise, Additive slope in the local level, and Additive seasonality (ETS-AAA). This model serves as a connection to both the MARSS univariate model, its most direct generalization, which does not yet incorporate consumer sentiment, and the bivariate MARSS model, in which we intend to incorporate consumer sentiment and its observed proxy. Indeed, as mentioned in the Introduction, we interpret the former as the additive slope at the local level and treat the latter as a second observed state in addition to the monthly sales volume. Note that we can meaningfully compare the first three models (ETS-AAA, ETS-ANA, and the univariate MARSS) only in terms of forecast-oriented accuracy metrics such as RMSSE and MASE. Information criteria such as log-likelihood, AIC, BIC, and AICc are not directly comparable across these models due to the fundamental structural difference between univariate and multivariate formulations.

Figure 3. Toyota Camry’s Monthly Consumer Sentiment Signals Proxies across ChatGPT, LLaMA, and CardiffNPL (2009 - 2022) - ETS-ANN Decomposition - Fitted Values and Residuals - Zero Lag Correlation Matrix



Source: Our analysis of data collected from Twitter

Nevertheless, the univariate MARSS model consistently outperforms the ETS models, indicating the advantages of state-space modeling even in a univariate setting. Among the models incorporating sentiment, the one based on ChatGPT shows the best in-sample fit but tends to flatten somewhat in the out-of-sample phase, closely tracking sales volume but possibly overfitting the training data.

In contrast, both LLaMA and CardiffNLP provide a more balanced fit between in-sample and out-of-sample sets. LLaMA performs slightly better in the in-sample period, whereas CardiffNLP outperforms LLaMA in the out-of-sample horizon. Notably, CardiffNLP yields the highest number of robust forecasting solutions, suggesting it may offer better generalization across different temporal segments.

Figure 4 compares the forecasting performance of the ETS-ANA model (top panel) and the univariate MARSS model (bottom panel) for the Toyota Camry’s monthly sales volumes. For the sake of clarity, we restrict the visualization to the period April 2018 -

December 2022. In both panels, the solid grey line represents the actual sales during the training period, while the black line indicates observations from the test set. The brown line shows the fitted values for the training period, and the magenta line represents the model point forecasts for the test period. The shaded areas of pale green, cyan, and orange indicate the 80%, 85%, and 90% prediction bands, respectively, enclosed between the green, blue, and red lines denoting the 80%, 85%, and 90% prediction intervals. The ETS-ANA model (top panel) produces forecasts with higher volatility and wider prediction bands during the test period, reflecting greater uncertainty in future sales projections.

*Table 3. Information criteria and accuracy measures<sup>3</sup> for the Toyota Camry's monthly sales volumes: ETS-ANA model, ETS-AAA model, MARSS univariate, and MARSS bivariate with ChatGPT proxy of consumer sentiment. Here, the letter y [resp. z] represents the Toyota Camry's monthly sales volume [resp. ChatGPT consumer sentiment proxy]*

Models	logLik	AIC	BIC	AIC <sub>c</sub>	RMSE	MAE	MAPE	SMAPE %	MASE	RMSSE
ETS-ANA (y)	-1591.434	3212.869	3257.623	3216.561	9440.055	8610.602	39.914	15.902	1.669	1.389
ETS-AAA (y)	-1592.126	3218.252	3268.973	3223.033	10809.330	10241.120	46.714	18.321	1.985	1.594
MARSS (y) No sent.	-1567.278	314.555	3184.319	3156.185	7291.802	6087.463	28.754	11.831	1.180	1.073
MARSS (y) ChatGPT	-1506.265	3036.531	3072.334	3038.876	3321.115	2816.744	12.627	6.231	0.546	0.488
MARSS (z) ChatGPT	-1506.265	3036.531	3072.334	3038.876	0.217	0.114	32.600	12.809	0.780	1.060
MARSS (y) LLaMA	-1844.953	3713.907	3749.710	3716.253	497.907	4042.257	19.105	8.465	0.784	0.732
MARSS (z) LLaMA	-1844.953	3713.907	3749.710	3716.253	0.161	0.104	50.148	20.349	0.679	0.796
MARSS (y) CardiffNPL	-1825.593	3675.186	3710.989	3677.532	3667.565	2735.730	12.197	5.741	0.530	0.539
MARSS (z) CardiffNPL	-1825.593	3675.186	3710.989	3677.532	0.148	0.060	87.304	5.741	0.453	0.864

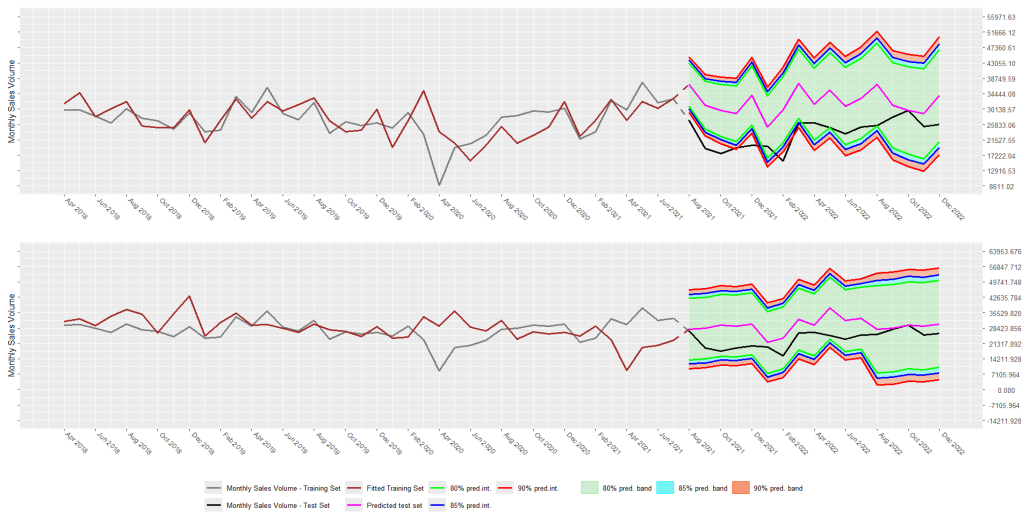
Fonte: nostre elaborazioni su dati estratti da Twitter

<sup>3</sup> *Information criteria:* AIC=Akaike information criterion, AIC<sub>c</sub>=Akaike information criterion corrected, BIC=Bayes Information Criterion (logLik=log-likelihood); *accuracy measures:* RMSE=Root Mean Square Error, MAE=Mean Absolute Error, MAPE=Mean Absolute Percentage Error, SMAPE=Symmetric mean absolute percentage error, MASE=Mean Absolute Scaled Error, RMSSE= Root Mean Square Scaled Error.

In contrast, the univariate MARSS model (bottom panel) demonstrates narrower prediction bands and smoother forecast trajectories, suggesting a more stable representation of sales dynamics. Overall, the ETS-ANA model is more sensitive to seasonal fluctuations, albeit at the cost of higher forecast uncertainty. Meanwhile, the MARSS univariate approach prioritizes stability over the ability to adapt to short-term variations.

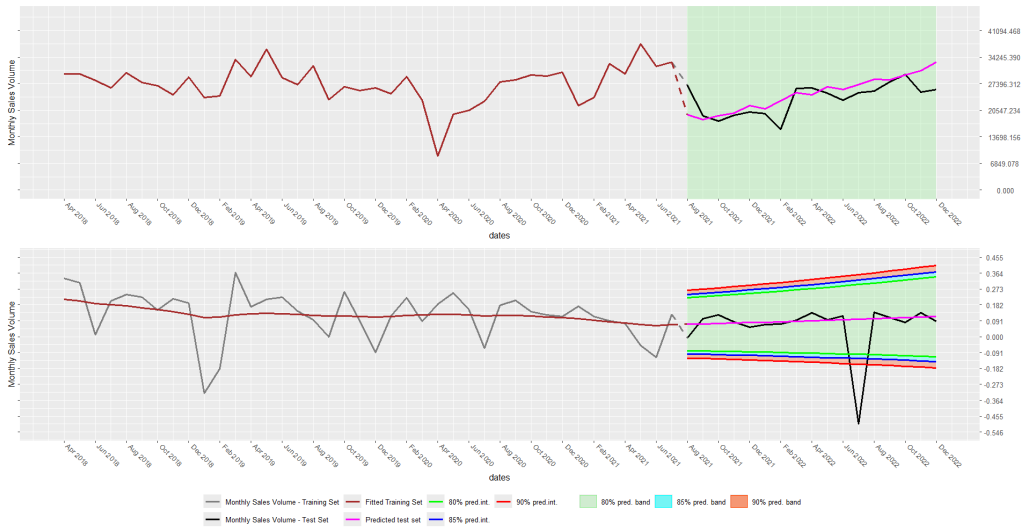
Figure 5 illustrates how our bivariate MARSS model, incorporating the CardiffN-PL sentiment mean, forecasts the Toyota Camry’s monthly sales volume for the same period, April 2018 - December 2022. This analysis explores the potential influence of public sentiment on sales forecasts. The plot retains the graphic features outlined for Figure 4. Referring to the above panel, due to the smoothing procedure of the Kalman filter, the forecasted training set of the monthly sales volume closely follows the actual sales volume, such that the brown line overlaps the grey line in the plot. The model’s high accuracy prevents it from fully representing the prediction bands, with only the 80% prediction band being visible.

Figure 4. ETS-ANA a univariate MARSS Forecast of the Toyota Camry’s Monthly Sales with Multiple Prediction Intervals



Fonte: nostre elaborazioni su dati estratti da Twitter

Figure 5. Bivariate MARSS model: predicted Toyota Camry's Monthly Sales Volume and CardiffNPL sentiment mean



Fonte: nostre elaborazioni su dati estratti da Twitter

The CardiffNPL sentiment average (bottom panel) exhibits a weak trend and resembles a LOESS curve, a consequence of the smoothing procedure. However, the predictive performance is generally good, except for the peak in April 2020, which coincided with the onset of the COVID-19 pandemic, suggesting a potential impact of external shocks on the model's accuracy.

## 5. Conclusions

Despite the increasing difficulty in accessing user review data due to tightening platform restrictions and growing concerns over data privacy, sentiment-based models continue to demonstrate strong predictive potential in forecasting consumer behavior and market dynamics. Given the increasing difficulty in obtaining high-quality user-generated reviews – due to changes in platform access policies and data availability constraints – refining the methodology for extracting sentiment signals has become essential. To this end, future work should prioritize a structured, multi-step approach aimed at enhancing the signal-to-noise ratio in sentiment data analysis. Specifically, we propose applying topic modeling techniques (e.g., BERTopic or neural topic models) to identify thematically coherent clusters within the corpus. This step enables the exclusion of irrelevant or noisy content (e.g., generic chatter, spam, or off-topic discussions) and the isolation of semantically meaningful segments such as complaints about

vehicle reliability or praise for driving experience. Once we identify relevant topics, we can conduct sentiment analysis within each topic cluster. This targeted approach ensures a more accurate alignment between sentiment scores and domain-specific concerns. By distinguishing sentiment related to, for example, “battery problems” versus “interior comfort”, the extracted emotional valence becomes both more interpretable and analytically precise. Subsequently, the sentiment scores should be aggregated at regular intervals (e.g., monthly) and smoothed using time series decomposition techniques (such as ETS or STL). This step further mitigates the impact of short-term volatility or noise, producing a more stable signal suitable for downstream forecasting models. Finally, we can train predictive models (e.g., MARSS or gradient boosting regressors) using these topic-filtered and temporally aggregated sentiment signals, and evaluate their performance using standard forecasting metrics such as RMSE, MAE, MASE, and RMSSE. Such a pipeline, spanning topic detection, sentiment filtering, and time series smoothing, is expected to yield sentiment signals that are not only more robust but also more predictive of real-world outcomes, such as sales volumes or customer satisfaction indicators.

## References

- BARBIERI, F., CAMACHO-COLLADOS, J., NEVES, L., & ESPINOSA-ANKE, L. (2020). TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. *Findings of the Association for Computational Linguistics: EMNLP 2020*, 1644-1650. <https://aclanthology.org/2020.findings-emnlp.148/>
- BARBIERI, F., CAMACHO-COLLADOS, J., ESPINOSA-ANKE, L., & NEVES, L. (2022a). TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. *Proceedings of the 2020 Findings of the Association for Computational Linguistics: EMNLP*, 1644-1650. <https://doi.org/10.18653/v1/2020.findings-emnlp.148>
- BARBIERI, F., ESPINOSA ANKE, L., & CAMACHO-COLLADOS, J. (2022b). *TweetEval: Unified benchmark and comparative evaluation for tweet classification*. *Natural Language Engineering*, 28(2), 289-326. <https://doi.org/10.1017/S1351324921000424>
- BENDER, E. M., GEBRU, T., McMILLAN-MAJOR, A., & SHMITCHELL, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT)* (pp. 610-623). <https://doi.org/10.1145/3442188.3445922>
- BOMMASANI, R., HUDSON, D. A., ADELI, E., ALTMAN, R., ARORA, S., VON ARX, S., & LIANG, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint*, arXiv:2108.07258
- BROWN, T. B., MANN, B., RYDER, N., SUBBIAH, M., KAPLAN, J., DHARIWAL, P., NEELAKANTAN, A.,

- SHYAM, P., SASTRY, G., ASKELL, A., AGARWAL, S., KRUEGER, G., HENIGHAN, T., CHILD, R., RAMESH, A., ZIEGLER, D. M., WU, J., WINTER, C., HESSE, C., AMODEI, D. (2020). Language Models are Few-Shot Learners. *ArXiv*. <https://arxiv.org/abs/2005.14165>
- CHOWDHERY, A., NARANG, S., DEVLIN, J., BOSMA, M., MISHRA, G., ROBERTS, A., BARHAM, P., CHUNG, H. W., SUTTON, C., GEHRMANN, S., SCHUH, P., SHI, K., TSVYASHCHENKO, S., MAYNEZ, J., RAO, A., BARNES, P., TAY, Y., SHAZEER, N., PRABHAKARAN, V., FIEDEL, N. (2022). PaLM: Scaling Language Modeling with Pathways. *ArXiv*. <https://arxiv.org/abs/2204.02311>.
- DEVLIN, J., CHANG, M. W., LEE, K., & TOUTANOVA, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 2019*, 4171-4186.
- HOLMES E. E, WARD E. J., SCHEUERELL M. D., WILLS K. (2024). MARSS: Multivariate Autoregressive. State-Space Modeling. R package version 3. <https://CRAN.R-project.org/package=MARSS>.
- HOLT, C. C. (1957). *Forecasting seasonals and trends by exponentially weighted averages* (O.N.R.Memorandum No. 52). Carnegie Institute of Technology, Pittsburgh, USA.
- HYNDMAN, R., KOEHLER, A., SNYDER, R., GROSE, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting* 18(3), 439-454.
- HUTTO, C. J., & GILBERT, E. (2014). *VADER: A parsimonious rule-based model for sentiment analysis of social media text*. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 8, No. 1, pp. 216-225). AAAI Press.
- IEZZI, D. F., & MONTE, R. (2024a). Social media effects on sales: Consumer sentiment in a state-space model. In G. Giordano & M. Misuraca (Eds.), *New Frontiers in Textual Data Analysis (Studies in Classification, Data Analysis, and Knowledge Organization)*, pp. 77-90). Springer Nature.
- IEZZI D. F., & MONTE R. (2024b). Building deep-learning consumers' sentiment signals for sales forecasting: a comparative study. In: (a cura di): Anne Dister Dominique Longrée, *Mots comptés, textes déchiffrés*. p. 419-428, LOUVAIN: Presses universitaires de Louvain, ISBN: 978-2-39061-471-5.
- IEZZI, D. F., & MONTE, R. (2025). E-Word of Mouth in Sales Volume Forecasting: Toyota Camry US case study. *Big Data Research*, 41, Article 100542. <https://doi.org/10.1016/j.bdr.2025.100542>
- META AI. (2024). Introducing LLareMA 3. Retrieved from <https://ai.meta.com/blog/meta-llama-3>
- OPENAI. (2023). GPT-4 technical report. *arXiv preprint*, arXiv:2303.08774.
- TOUVRON, H., LAVRIL, T., IZACARD, G., MARTINET, X., LACHAUX, M., LACROIX, T., ROZIÈRE, B., GOYAL, N., HAMBRO, E., AZHAR, F., RODRIGUEZ, A., JOULIN, A., GRAVE, E., & LAMPLE, G. (2023). LLaMA: Open and Efficient Foundation Language Models, 2023b. *ArXiv*. <https://>

[arxiv.org/abs/2302.13971](https://arxiv.org/abs/2302.13971).

TOUVRON, H., MARTIN, L., STONE, K., ALBERT, P., ALMAHAIRI, A., BABAEI, Y., & SCIALOM, T. (2023). Llama 2: Open foundation and fine-tuned chat models, 2023b. *ArXiv*. <https://arxiv.org/abs/2307.09288>, 38.

VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, Ł., & POLSUKHIN, I. (2023). *Attention is all you need*. *arXiv*. <https://arxiv.org/abs/1706.03762>.



## ANALISI DELLA MOBILITÀ E DEL TURISMO: DATI PASSIVI, MACHINE LEARNING E LLM

### MOBILITY AND TOURISM ANALYTICS: PASSIVE DATA, MACHINE LEARNING, AND LLMs

*Bruno Zamengo<sup>1</sup>, Glauco Mantegari<sup>2</sup>, Denis Cappellari<sup>3</sup>*

#### Sommario

Il contributo propone una visione integrata di come dati di mobilità e turismo su larga scala – derivati da operatori di telefonia mobile (MNO), tracce GPS e Floating Car Data (FCD), integrati con sensori stradali e dati di ticketing – possano affiancare le statistiche ufficiali con misure più tempestive e granulari. Attraverso casi d’uso in ambito turistico, aeronautico e di mobilità generale, mostra come dati anonimizzati consentano di stimare presenze e flussi, pattern di spostamento, congestione e utilizzo delle infrastrutture, evidenziando vincoli di rappresentatività e requisiti di privacy. Discute approcci operativi per riconciliare sorgenti eterogenee e diverse granularità (data fusion e machine learning) e descrive il potenziale dei Large Language Models (LLM) per interagire in linguaggio naturale con dati strutturati e documenti, riducendo la dipendenza da dashboard e da interrogazioni SQL. Nel complesso, l’integrazione multi-sorgente con ML e interfacce GenAI human-centriche può aumentare l’accessibilità degli insight per la pianificazione dei trasporti, la gestione aeroportuale e il turismo sostenibile, mantenendo la valutazione di qualità e accuratezza.

#### Abstract

*This paper presents how large-scale mobility and tourism data – sourced from mobile network operators, GPS traces, Floating Car Data (FCD), traffic sensors, and ticketing systems – can complement official statistics with more timely, granular measures. Through applications in tourism, civil aviation, and general mobility, we show how anonymized data enable measurement of presences, travel patterns, congestion, and infrastructure usage, while requiring treatment of representativeness and privacy. We outline practical ways to reconcile heterogeneous sources and granularities (data fusion and machine learning) and discuss how Large Language Models (LLMs) can enable natural-language access to insights over structured and*

<sup>1</sup> Motion Analytica, Udine, Italy - e-mail: bruno.zamengo@motionanalytica.com (corresponding author).

<sup>2</sup> Motion Analytica, Udine, Italy - e-mail: glauco.mantegari@motionanalytica.com

<sup>3</sup> Motion Analytica, Udine, Italy - e-mail: denis.cappellari@motionanalytica.com

*unstructured assets, reducing reliance on SQL and dashboards. Use cases highlight operational value for sustainable tourism, transport planning, and airport management, alongside challenges in data quality and evaluation.*

## **Introduction**

Today, vast amounts of data from mobile operators, Floating Car Data (FCD), traffic loops, public transport ticketing, and GPS apps generate valuable insights for decision-makers in urban planning, tourism, and sustainable mobility.

This paper argues that passively generated big data – especially mobile network operator (MNO) records, GPS traces, and Floating Car Data (FCD) – can complement (rather than replace) official statistics by providing more timely and granular measurements, provided that privacy, sampling bias, and representativeness are explicitly addressed. Through three applications (tourism, civil aviation, and general mobility), we show how multi-source integration – via calibration, data fusion, and machine learning – yields actionable indicators for public and private decision-makers. We also discuss how Large Language Models (LLMs) can sit on top of curated datasets as a transversal interface, enabling natural-language access to validated insights beyond dashboards.

The paper is organised as follows: Section 1 focuses on tourism, Section 2 on civil aviation, Section 3 on general mobility, and Section 4 on data fusion, machine learning, and generative AI (including LLM-based assistants). The final section summarises the conclusions.

Across these domains, we treat mobility as a family of related phenomena that require different observational lenses. Our core claim is that deliberate multi-source integration (calibration, data fusion, and ML) combines complementary and non-redundant information, yielding decision-relevant indicators that remain unattainable when each source is analysed separately.

### **1. Tourism**

We explored mobile big data for tourism as a complement to traditional sources, providing more timely and granular indicators while requiring calibration and representativeness checks. Prior work shows how telecom data can quantify tourism pressure and externalities (e.g., Baldin *et al.*, 2024; Confente *et al.*, 2024). Methodological frameworks and public–private collaborations have also been proposed to integrate mobile positioning data into official tourism statistics in a privacy-compliant way (Saluveer *et al.*, 2020; United Nations Committee of Experts on Big Data and Data Science for Official Statistics, 2021; Cavallo *et al.*, 2022).

A study based on data provided by Motion Analytica and Vodafone Business Italia (Scotti *et al.*, 2024) analyses drivers of overnight stays and same-day visits in Lombardy (Italy). Compared with official statistics, mobile network data offers greater behavioural granularity but still requires calibration when scaling a single-operator sample (Fastweb–Vodafone Italia customer plus roamers) to population-level indicators.

Key considerations include uneven network infrastructure (affecting localisation precision) and market-penetration bias. We therefore validate volumes and seasonality against ISTAT/official counts at municipality and province level and release results only as privacy-preserving aggregates subject to minimum-count thresholds. Indicators are used to study determinants of stay duration and same-day visits, highlighting the role of events and intermodal nodes in shaping day-tripper behaviour and supporting targeted destination policies (Scotti *et al.*, 2024).

A second project by Motion Analytica, in collaboration with Vodafone Business Italia (now Fastweb–Vodafone Italia), analysed visitor flows in Gran Paradiso (PNGP) and in Abruzzo, Lazio and Molise (PNALM) National Parks to complement official statistics and support evidence-based management. Mobile phone data profiled daily and overnight visits (e.g., nationality and residence), and a second phase explored qualitative context via selected Meta hashtags (Scotti *et al.*, 2024).

Additional examples are reported in Zamengo *et al.* (2024), which document recurring patterns that can be directly translated into management-relevant indicators. Tourist presence peaks in summer – especially in August – and exhibits clear weekend spikes. The time series also captures the structural break induced by COVID-19: in 2020, foreign arrivals dropped sharply, while domestic tourism partly compensated for the decline. Cross-destination comparisons further show distinct visitor compositions, with PNALM attracting a larger share of Italian visitors and PNGP featuring a higher proportion of foreigners. Taken together, these indicators provide an empirical basis for destination management measures such as mitigating private-vehicle traffic, improving visitor experience through flow governance, and reducing pressure on sensitive ecosystems.

## 2. Aviation

The world of civil aviation includes two main types of actors: airports and airlines. For an airport, a strategic factor is its centrality within the network of air connections, which determines its ability to attract the highest number of travellers. This centrality can be assessed from two perspectives: inbound traffic analysis, referring to the airport's capacity to serve as an entry point for travellers heading to its surrounding region; and outbound traffic analysis, referring to its role as a gateway to remote destinations of interest for locals. In this context, mobile phone data (cross-referenced and enriched

with other data sources such as timetables) has proven to be a valuable resource (Motion Analytica, n.d.), especially when compared to other market solutions or official sources (Associazione Italiana Gestori Aeroporti, n.d.; Eurostat, n.d.). Such data provides airports with tools to estimate their catchment and distribution areas based on observed behaviour.

Basic models were initially complemented by an unsupervised study using clustering techniques, which gradually evolved into more sophisticated machine-learning solutions. A key enhancement involved the adoption of a graph-based data model to represent the global airport network, where nodes correspond to airports and edges represent direct flight connections weighted by flight duration. This structure enabled the modelling of mobility patterns across the air transport system, allowing for the identification of latent structures in passenger flows and the detection of anomalous or emergent behaviours. Building on this foundation, mobile phone data was integrated with actual flight departure and arrival times at airports, significantly enhancing the temporal precision of mobility estimates. To ensure that the insights derived from the sample population (primarily Fastweb–Vodafone Italia customers and roaming users) could be generalised to the entire population of air travellers, we incorporated official statistics from a preceding period. These statistics, however, posed two main challenges: they were aggregated at a less granular level than our model estimates, and they were typically published with a significant delay.

We infer probabilistic flight-level passenger assignments by aligning airport dwell episodes derived from mobile network data with scheduled and actual flight events; results are then aggregated to route/airport/time bins. The statistical unit is the passenger itinerary, counted at origin–destination (O-D) level (first origin and final destination), while preserving the possibility to analyse intermediate legs when needed. Basic models were initially complemented by an unsupervised phase (e.g., clustering of mobility signatures), which then evolved into more structured machine-learning solutions for passenger–flight linkage and for reconciling high-resolution mobile estimates with coarser official aggregates.

To address the former issue, we employed data fusion techniques to reconcile the differing levels of granularity, aligning high-resolution mobile data with coarser official aggregates. For the latter challenge, we developed and applied time-predictive algorithms capable of forecasting official statistics based on real-time behavioural signals. This dual approach enabled us to produce timely, high-resolution estimates of air travel behaviour across all major Italian airports, over forty in total, extending well beyond the limitations of any single data source: the output consists of an integrated analytical model that describes users' aviation experience, from home location to tourism destination, with flight-level detail. The data can be explored across different levels of aggregation.

In the following paragraphs, we present details on the estimation precision and an application of the analytical results.

To validate the robustness of our estimates, we conducted precision assessments and comparative analyses against official data sources. In the following paragraphs, we present key findings from this process, along with selected applications of the analytical results. Figure 1 reports the comparison between our estimates and official data for routes (Eurostat, n.d.), showing that the overall correlation per route exceeds 99% for domestic flights and 95% for international flights, with linear coefficients of 1.00 and 1.04, respectively (official data downloaded in August 2022). Figure 2 also presents the Mean Absolute Error (MAE) index (Morley *et al.*, 2018) which, as expected, shows that the accuracy of the estimates increases with the size of the phenomenon: routes with higher passenger traffic report significantly lower MAE values. Consequently, for smaller-scale phenomena with greater variability, we typically aggregate the data into statistics that describe larger phenomena, without introducing additional heuristics or artificially adjusting the data.

*Figure 1. The x-axis shows the number of passengers recorded by official statistics, while the y-axis shows our big data estimates. Source: Authors' elaboration on Motion Analytica data*

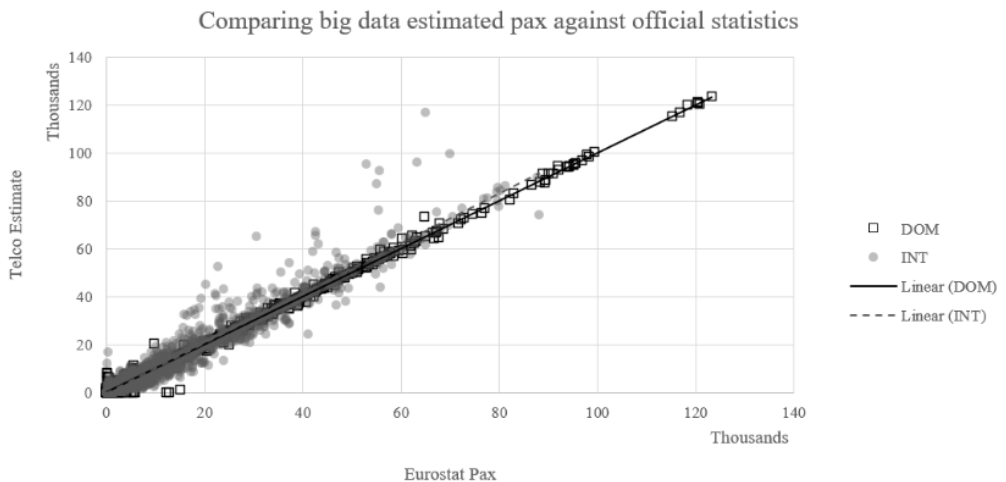
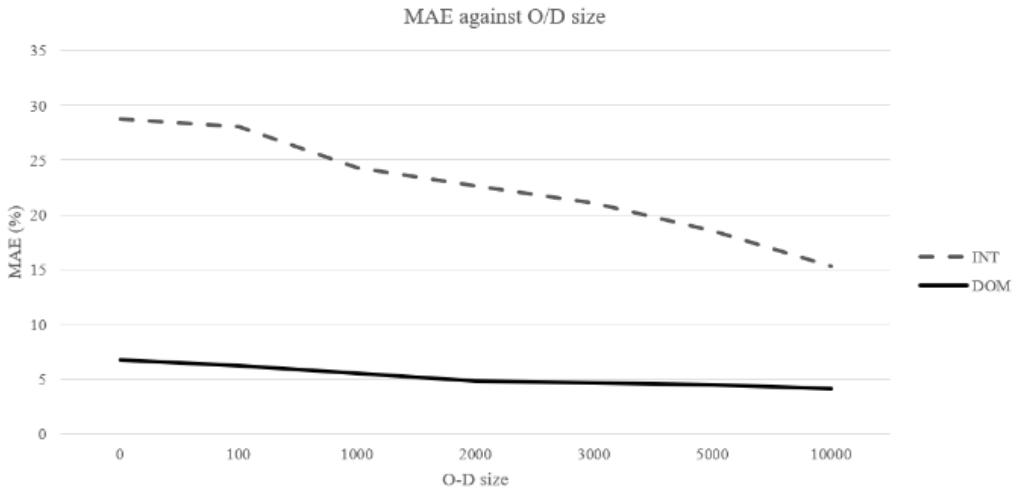


Figure 2. Origin–destination (O–D) pairs are grouped into bins based on monthly passenger volume. The “O–D size label” denotes the bin’s lower bound (e.g., “100” corresponds to 100–999 passengers per month). The dashed line shows the MAE index for international O–Ds, while the continuous line refers to domestic O–Ds. Source: Authors’ elaboration on Motion Analytica data



One of the added values of big data estimates compared to official statistics is the ability to explore additional details. In this case, for example, it is possible to analyse route-level data based on users’ estimated habitual residence inferred from their mobility behaviour.

The maps in Figure 3 show the market share of the airports of Bologna, Firenze, and Pisa broken down by the municipality of residence of travellers. The first map covers all international routes, while the latter shows routes to Germany. For privacy reasons, only municipalities with a minimum number of travellers are shown.

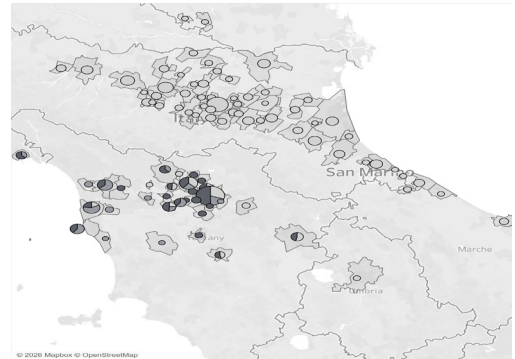
The ability to determine the origin of travellers from a particular airport or even a specific route can be used to characterise an airport’s catchment area based on observed behaviour. This approach also helps quantify phenomena such as leakage, where travellers choose to use an airport farther from their home (e.g. due to the availability of routes not offered at nearby airports). For instance, Figure 3 shows that Bologna airport tends to exhibit the broadest catchment across municipalities. Florence and Pisa display an interesting pattern: while Pisa generally appears to draw travellers from a wider area, Florence is higher than it when focusing on connections to Germany (right map).

Figure 3. Municipality-level maps of air traveller market shares among Bologna (light grey), Florence (dark grey) and Pisa (mid grey). Pie charts show airport shares by residence. Source: Authors' elaboration on Motion Analytica data. Note: Axis values or scales are not shown, as the underlying data is confidential

International Air Travellers Share of Competitors



Air Travellers to Germany Share of Competitors



### 3. Big data for general mobility

The study of mobility through big data and ML/AI accelerated significantly during the COVID-19 pandemic, when numerous studies leveraged GPS- and mobile-network-based signals to monitor behavioural changes at scale (e.g., DePhillipo *et al.*, 2021; Jewell *et al.*, 2021; Xiong *et al.*, 2020). This line of work builds on earlier research using MNO data to estimate mobility demand and origin–destination flows (Calabrese *et al.*, 2011, 2014), and motivates multi-source strategies that combine complementary data strengths.

Beyond pandemic applications, MNO data has long been used to measure mobility and to build Origin–Destination matrices on broad samples (Calabrese *et al.*, 2011, 2014). GPS data can be biased because it is generated by specific user segments and depends on opt-in apps and sampling settings; consequently, robust strategies often combine insights from both MNO and GPS sources, leveraging their respective strengths. In addition to data that detect the movement of people, others focus on vehicles: Floating Car Data (FCD) from vehicle black boxes has been used to analyse parking demand and traffic congestion (Mannini *et al.*, 2018; Nalubega, 2023; He *et al.*, 2016).

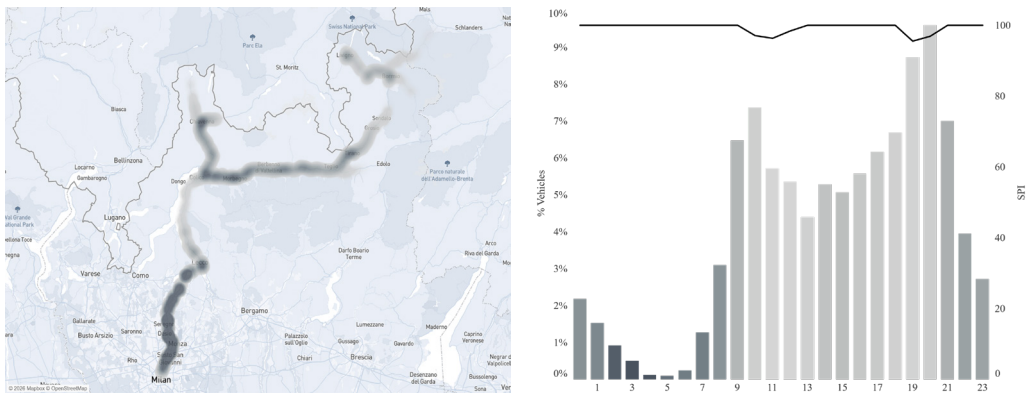
Floating Car Data (FCD) is passively generated trace data collected from equipped vehicles (e.g., black boxes, OEM or insurer fleets), shifting the focus from “people mobility” to “vehicle mobility”. Compared with mobile network and app-based GPS data – which can be sparse, self-selected, or limited in available attributes – FCD typically provide higher-frequency vehicle trajectories enriched with variables such as

speed (and, when available, vehicle class/engine information), enabling additional use cases in traffic operations and infrastructure monitoring. In this section, we present two applications: the former estimates traffic trends on road segments and their evolution over time (e.g., recurring congestion windows); the latter uses FCD to estimate usage of airport parking facilities and competing parking lots in the vicinity. As FCD coverage depends on provider penetration and fleet composition, results are interpreted as robust indicators of trends and segment ranking and, where possible, cross-checked against conventional sensors.

A common application of FCD is traffic analysis, extending traditional sensor-based approaches (e.g., inductive loops). While conventional sensors enable various analyses, such as vehicle counts or local speed measurements, FCD offers a more comprehensive perspective by tracking vehicles along their entire routes. This enables detailed insights segmented by vehicle class, manufacturer (make), model, and more, or by filtering for specific vehicle types. In this application, for example, we analysed only vehicles that travelled at least 500 meters along the “Olympic Way,” the route that will connect Milano with the cities hosting selected events of the Milano-Cortina 2026 Winter Olympics.

We measured both the number of vehicles and their speeds across each segment of the Olympic Way; results are aggregated at road-segment/hour level and individual trajectories are not released. Figure 4 shows, on the left, the distribution of vehicle counts along the entire Olympic Way and, on the right, the hourly vehicle counts and Speed Performance Index (SPI) for the same segment.

Figure 4. FCD-based traffic analysis along the “Olympic Way” (Milano-Cortina 2026 corridor). Source: Authors’ elaboration on Motion Analytica data



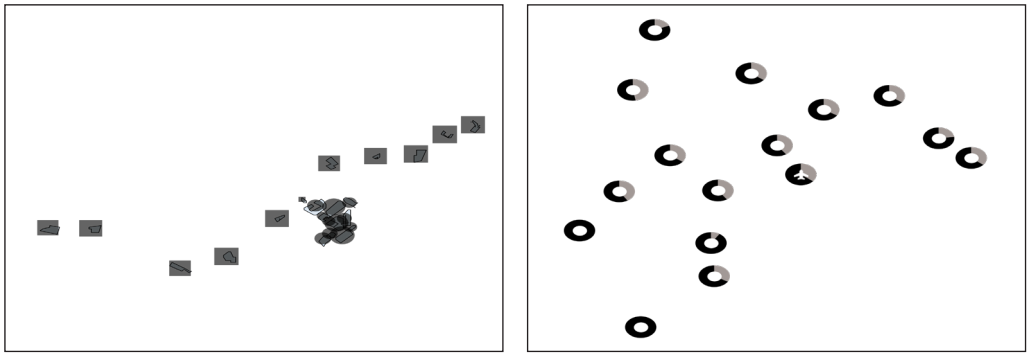
4.a. FCD traffic heatmap: more intense colour indicates higher vehicle counts per road segment

4.b. Hourly vehicle counts (bars) and Speed Performance Index, SPI (line) on the most trafficked segment

Overall, the analysis shows that the most heavily used sections of the route are between Milan and Lecco. On the busiest segment, the estimated speed distribution suggests that a non-negligible share of vehicles (approximately 10%) exceeds the posted limit. The hourly profile also highlights evening peak demand (hours 19-21), which coincides with the lowest SPI values, signalling increased congestion during peak usage.

These examples illustrate how Floating Car Data (FCD) can provide stakeholders with actionable insights into road usage, congestion patterns, and driver behaviour, supporting more informed decisions on traffic management policies and infrastructure investments.

*Figure 5. Floating car data used to analyse parking usage around an airport*



*5.a. Parking usage: airport (circles) vs competitor parking (squares), by volume of parked vehicles.*

*5.b. Vehicle origin/destination: donuts compare airport parking (light grey) vs competitor parking (black) by province (airplane = airport's province).*

Source: Authors' elaboration on Motion Analytica data. Note: Axis values or scales are not shown, as the underlying data is confidential

The analysis of airport parking usage is relatively straightforward when focusing on airport-owned parking facilities, as it is reasonable to assume that vehicles parked in these areas belong to travellers or individuals directly connected to airport-related activities. However, analysing nearby third-party parking lots is more complex, as users may be visiting other facilities in the area rather than the airport itself. To address this, we implemented a selection strategy based on the vehicle's origin (or destination) before (or after) the parking session, as well as the session duration. We specifically focused on stops whose duration is compatible with at least one round-trip flight from the selected airport. This analysis can help the airport understand its parking leakage, that is, how many vehicles prefer competitor parking rather than airport parking. The analysis

can be drilled down in different ways, such as by vehicle origin or engine type, as shown in Figure 5.

In the studies by Costa and Silvestri (2021) and Cavattoni, Comin, and Silvestri (2024), shared mobility data, such as bicycles and scooters, were used to optimise fleet distribution and to measure the impact of COVID-19 on shared mobility. In the case of FCD and shared mobility data, coverage is tied to the provider's market penetration and, since these data pertain to vehicles, they cannot be directly linked to other sources focused on people. However, integrating multiple sources – while accounting for their limitations and unique characteristics – can provide a more complete picture of the phenomenon.

As in other sectors, the use of big data for mobility and the insights derived from it can serve various purposes. Two major Italian institutions (Ministero delle Infrastrutture e della Mobilità Sostenibili, 2022; Tartaglia et al., 2024) identify the following main applications: (i) measure the real demand for mobility comprehensively, detailing it by means of transportation; (ii) plan and improve public transport services and infrastructure; (iii) plan the ecological transition to support decisions that promote the shift towards more sustainable mobility; and (iv) monitor the evolution of mobility behaviours over time and measure the effectiveness of chosen strategies and implemented interventions.

These data have also been used to explore the potential for supplementing official statistics. For example, in our joint work with ISTAT and Vodafone (Radini et al., 2021) we studied the population habitually residing in an area and constructed commuting matrices. From these experiments and other initiatives by official statistical agencies, several European projects have emerged with the aim of defining a common framework for all member countries. This framework seeks to integrate traditional sources with new data sources in the production of official statistics for the European Community (Ricciato & Coluccia, 2023; Ricciato *et al.*, 2020; Ricciato, 2018).

*Table 1. Cross-case synthesis: why multi-source integration matters*

Domain	Sources	Output and decision use	Why multi-source
Tourism (1)	MNO + official tourism stats (calibration); optional social/UGC for qualitative context	Timely indicators of presences/flows and day-tripper behaviour → destagionalisation and flow governance	Official stats: coarse/delayed. MNO: sampling/market bias and localisation uncertainty. Fusion yields calibrated aggregates.
Aviation (2)	MNO + flight ops/ timetables + official stats	Catchment/leakage and O-D with flight-level detail → airport and airline strategy	MNO: limited flight attribution. Schedules: no profiling. Official stats: coarse/delayed. Fusion aligns detail to trusted totals.
General mobility (3)	MNO + map/POI; FCD + sensors + map	Traffic trends, congestion/speeds, parking usage → transport planning and traffic policies	MNO: broad but coarse. FCD: operational but biased/low penetration. Sensors: local only. Fusion combines scale and dynamics.

Overall, the cases highlight recurring trade-offs: MNO data maximise coverage and timeliness but require calibration for bias and localisation uncertainty; GPS traces add behavioural detail but are self-selected; FCD provides operational dynamics but depends on penetration. Data fusion and ML bridge these regimes by reconciling granularities and transferring information from higher-quality labels (e.g., official aggregates) to larger-scale passive signals.

#### 4. ML, AI & GenAI

Zheng (2015) presents a literature review in the field of “data fusion”, analysing various use cases and their respective resolution methodologies. Mobility (and tourism) data can benefit from these techniques because they originate from different sources and have varying granularities and available features. Motion Analytica is developing a data fusion technique for MNO and GPS data using ML models, one of the techniques indicated in Zheng (2015).

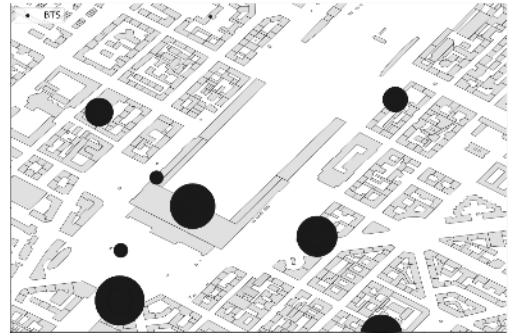
The process begins by transforming MNO data (originally georeferenced to cell towers) into sequences of point-based positions using a proprietary algorithm known as pseudo-GPS (p-GPS). In parallel, a GPS dataset enriched with POI information from Overture Maps Foundation (n.d.) is used to train a model that infers the most likely buildings in which users are located. By applying this trained model to the p-GPS data, we generate new geospatial data that closely resemble real GPS traces, but is derived from a much larger user base, as it originates from the telecom’s operator customer network. Figure 6 summarizes the workflow and its outputs (A–F).

The mapping model is trained on opt-in users with both GPS and MNO signals; performance may degrade under domain shift (different urban morphology, antenna density, handset behaviour), so calibration is required when transferring across areas/operators.

Figure 6. Supervised fusion of MNO and GPS data: workflow and mapping outputs



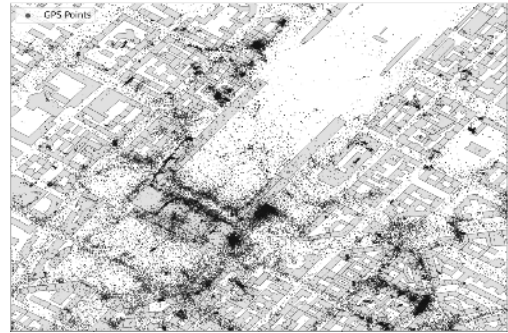
A. Overture buildings/metadata (e.g., commercial vs residential).



B. User presence by antenna (bubble size  $\propto$  connected users).



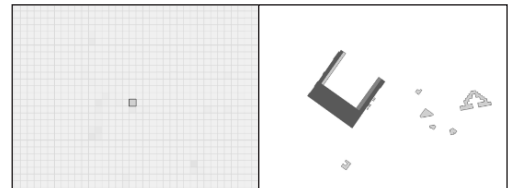
C. p-GPS output from MNO data (artefacts/clustering visible).



D. GPS points for the same users (ground truth).



E. Building  $\rightarrow$  pixel weights (darker = higher likelihood).



F. Pixel  $\rightarrow$  building weights (darker = higher probability).

Source: Authors' elaboration on Motion Analytica data

In November 2022, ChatGPT was announced and released to the public, creating a new wave of interest in AI and ML. However, several works utilizing these types of technologies had already been in use for some time. For example, the solution presented in Section 3 also employs standard ML techniques to accurately associate a passenger with a flight. A more extensively documented case can be found in Kim *et al.* (2022) (and further explored in its bibliography), where the authors propose a machine learning model that classifies modes of travel based on movement patterns detected via GPS data, using a convolutional neural network.

Recent advances in generative AI and Large Language Models (LLMs) enable natural-language access to data and documents. In our work (Padoan *et al.*, 2024), we explored LLM-based assistants that translate user questions into retrieval and query actions over curated assets, reducing reliance on dashboards and manual SQL while making insights accessible to a broader audience.

In our experiments, providing domain-specific context over curated structured and unstructured assets enabled an LLM-based assistant to answer mobility questions and to trigger controlled retrieval/query actions, delivering both descriptive statistics and selected analytical outputs.

The development and use of this type of chatbot, based on agents and large language models (LLMs), introduces several challenges, one of the most critical being the evaluation of its accuracy.

LLM-based agents introduce operational risks (e.g., hallucinations), so outputs should be grounded in deterministic queries over curated, privacy-preserving aggregates, with access controlled through permissions and auditable logs. In practice, the LLM is best used to translate natural-language questions into constrained actions (e.g., read-only SQL or retrieval), while numeric results are computed from query outputs rather than free-text generation.

In an update to our previous work, presented at the GenAI4MoDA 2025 conference, we addressed the evaluation challenge by proposing an approach inspired by common practices in big data and machine learning evaluation (Fantin *et al.*, 2025). The process steps are:

1. We created a set of training questions for the chatbot. For each question, we provided the correct SQL query via prompting, guiding the chatbot toward the expected logic.
2. We then built a second set of questions, each paired with a manually written and verified SQL statement, referred to as the gold query.
3. We repeatedly sampled questions at random from either set (with replacement), submitting them to the chatbot.
4. For each interaction, we saved the SQL statement generated by the chatbot.

5. Finally, we compared the dataset returned by the chatbot-generated query with the one produced by the gold query, measuring either perfect matches or degrees of mismatch.

This evaluation framework allowed us to quantify the accuracy of the “chat with data” module. It also provided actionable insights into which types of questions require improved prompting strategies to enhance the chatbot’s performance.

To evaluate the “chat with data” module, we track whether generated SQL executes successfully and compare the resulting outputs against gold queries using exact-match and relaxed equivalence criteria (e.g., tolerances for numeric aggregates and overlap for rankings). When mismatches occur, we log error categories (e.g., wrong filter/join/granularity), which supports targeted prompt and tooling improvements.

## Conclusions

This work shows how large-scale mobility and tourism data – sourced from mobile networks, GPS, Floating Car Data (FCD), and shared mobility systems – can complement traditional statistics with granular, timely, actionable insights for public and private stakeholders.

Across the case studies, multi-source integration and ML help reconcile coverage, granularity, and accuracy trade-offs, producing indicators that support transport planning, airport strategy, and sustainable tourism. Within this framing, LLMs are not an alternative modelling approach but a transversal interface layer that helps non-specialists access curated indicators and documentation through natural language, while keeping outputs grounded in validated aggregates and explicit quality checks.

These approaches support decision-making in transport planning, airport operations, sustainable tourism, and infrastructure policy. Challenges around heterogeneity, privacy, representativeness, and evaluation remain, but multi-source integration – paired with intelligent automation and human-centric interfaces – will continue to shape mobility and tourism analytics. Among the most promising directions are: extending data fusion techniques and calibration protocols and improving LLM robustness on complex analytical queries through structured evaluation pipelines.

## Disclosure

The authors declare no conflicts of interest. Generative AI tools were used to improve the readability of the manuscript; ideas, analyses, and conclusions remain the sole responsibility of the authors.

## References

- ASSOCIAZIONE ITALIANA GESTORI AEROPORTI. (n.d.). *Tabella dati di traffico*. Retrieved from <https://assaeroporti.com>
- BALDIN, A., BERTOCCHI, D., & CAMATTI, N. (2024). Crowding-in and (temporary) crowding-out in Venice: The effect of cultural events on residents. *Tourism Economics*, 31(5), 984-1005
- BERTOCCHI, D., CAMATTI, N., & VAN DER BORG, J. (2021). Tourism peaks on the Three Peaks: Using big data to monitor where, when and how many visitors impact the Dolomites UNESCO World Heritage Site. *Rivista Geografica Italiana*, 3. <https://doi.org/10.3280/rgioa3-2021oa12532>
- CALABRESE, F., FERRARI, L., & BLONDEL, V. D. (2014). Urban sensing using mobile phone network data: A survey of research. *ACM Computing Surveys*, 47(2), Article 25. <https://doi.org/10.1145/2655691>
- CALABRESE, F., DI LORENZO, G., LIU, L., & RATTI, C. (2011). Estimating origin-destination flows using mobile phone location data. *IEEE Pervasive Computing*, 10(4), 36-44. <https://doi.org/10.1109/MPRV.2011.41>
- CALABRESE, F., COBELLI, E., FERRAIUOLO, V., MISSERI, G., PINELLI, F., & RODRIGUEZ, D. (2021). Using Vodafone mobile phone network data to provide insights into citizens' mobility in Italy during the Coronavirus outbreak. *Data & Policy*, 3, e22. <https://doi.org/10.1017/dap.2021.18>
- CAVATTONI, M., COMIN, M., & SILVESTRI, F. (2024). COVID-19 pandemic's enduring impact on urban mobility: The case of free-floating bike sharing in Padova, Italy. *Transportation Research Record: Journal of the Transportation Research Board*, 2678(8), 1046–1058. <https://doi.org/10.1177/03611981231223977>
- CAVALLO, L., CERASTI, E., DI TORRICE, M., RIGHI, A., SANTORO, M. T., TUOTO, T., VALENTINO, L., DI SORTE, D., ROSSI, M., ZARAMELLA, A., BERTOCCHI, D., MANTEGARI, G., & ZAMENGO, B. (2022). Exploring mobile network data for tourism statistics: The collaboration between Istat and Vodafone Business Italia. *Rivista di Statistica Ufficiale*, 3, 43-76. [https://doi.org/10.1481/ISTATRIVISTASTATISTICAUFFICIALE\\_3.2022.02](https://doi.org/10.1481/ISTATRIVISTASTATISTICAUFFICIALE_3.2022.02)
- CONFENTE, I., MAZZOLI, V., CAMATTI, N., & BERTOCCHI, D. (2024). Integrating tourists' walk and talk: A methodological approach for tracking and analysing tourists' real behaviours for more sustainable destinations. *Journal of Sustainable Tourism*, 32(11), 2323-2343. <https://doi.org/10.1080/09669582.2024.2322120>
- COSTA, E., & SILVESTRI, F. (2021). On the bike spreading problem. *arXiv* (arXiv:2107.00761). <https://doi.org/10.48550/arXiv.2107.00761>
- DEPHILLIPO, N. N., CHAHLA, J., BUSLER, M., & LAPRADE, R. F. (2021). Mobile phone GPS data and prevalence of COVID-19 infections: Quantifying parameters of social distancing in the U.S. *Archives of Bone and Joint Surgery*, 9(2), 217-223. <https://doi.org/10.22038/abjs.2020.48515.2404>
- EUROSTAT. (n.d.). *Air transport measurement – passengers and freight*. Retrieved from <https://>

ec.europa.eu/eurostat/

- FANTIN, L., ANTONELLI, M., CESETTI, M., IRTO, D., ZAMENGO, B., & SILVESTRI, F. (2025). Design and testing of an agent chatbot supporting decision making with public transport data. *arXiv* (arXiv:2505.22698). Retrieved from <https://arxiv.org/abs/2505.22698>
- HE, F., YAN, X., LIU, Y., & MA, L. (2016). A traffic congestion assessment method for urban road networks based on Speed Performance Index. *Procedia Engineering*, 137, 425-433. <https://doi.org/10.1016/j.proeng.2016.01.277>
- JEWELL, S., FUTOMA, J., HANNAH, L., MILLER, A. C., FOTI, N. J., & FOX, E. B. (2021). It's complicated: Characterizing the time-varying relationship between cell phone mobility and COVID-19 spread in the U.S. *npj Digital Medicine*, 4, 152. <https://doi.org/10.1038/s41746-021-00523-3>
- KIM, J., KIM, J. H., & LEE, G. (2022). GPS data-based mobility mode inference model using long-term recurrent convolutional networks. *Transportation Research Part C: Emerging Technologies*, 135, 103523. <https://doi.org/10.1016/j.trc.2021.103523>
- MANNINI, L., CIPRIANI, E., CRISALLI, U., & GEMMA, A. (2018). FCD data for on-street parking search time estimation. *IET Intelligent Transport Systems*, 12, 664-672. <https://doi.org/10.1049/iet-its.2017.0223>
- MINISTERO DELLE INFRASTRUTTURE E DELLA MOBILITÀ SOSTENIBILI. (2022). *Mobilità e logistica sostenibili – Analisi e indirizzi strategici per il futuro*. Retrieved from <https://www.mit.gov.it/comunicazione/news/mobilita-e-logistica-sostenibili-pubblicato-il-documento-strategico-del-mimss>
- MORLEY, S. K., BRITO, T. V., & WELLING, D. T. (2018). Measures of model performance based on the log accuracy ratio. *Space Weather*, 16, 69-88. <https://doi.org/10.1002/2017SW001669>
- MOTION ANALYTICA. (n.d.). *Analysis of the catchment area of Milan airports*. Retrieved from <https://www.motionanalytica.com/en/portfolio/sea-milan-airports-case-study-en/>
- NALUBEGA, S. I. (2023). Development of a Floating Car Data (FCD) model to evaluate traffic congestion: A case of Kampala, Uganda (Unpublished master's thesis). Stellenbosch University, Stellenbosch, South Africa.
- NALIN, A., VIGNALI, V., LANTIERI, C., CAPPELLARI, D., ZAMENGO, B., & SIMONE, A. (2024). Assessing veracity of big data: An in-depth evaluation process from the comparison of mobile phone traces and ground truth data in traffic monitoring. *Journal of Transport Geography*, 118, 103930. <https://doi.org/10.1016/j.jtrangeo.2024.103930>
- NURMI, O., & PIELA, P. (2019). The use of mobile phone data in tourism statistics. In *Proceedings of the 62nd ISI World Statistics Congress* (Special Topic Session, Vol. 4, pp. 135-142). Retrieved from [https://2019.isiproceedings.org/Files/6.Special-Topic-Session\(STS\)-Volume-4.pdf](https://2019.isiproceedings.org/Files/6.Special-Topic-Session(STS)-Volume-4.pdf)
- OVERTURE MAPS FOUNDATION. (n.d.). *Overture Maps*. Retrieved from <https://overturemaps.org>
- PADOAN, L., CESETTI, M., BRUNELLO, L., ANTONELLI, M., ZAMENGO, B., & SILVESTRI, F. (2024).

- Mobility ChatBot: Supporting decision making in mobility data with chatbots. In *2024 25th IEEE International Conference on Mobile Data Management (MDM)* (pp. 295-300). <https://doi.org/10.1109/MDM61037.2024.00061>
- PANAROTTO, A., CATTELAN, M., & BELLIO, R. (2024). State-space models for clustering of compositional trajectories. In J. Einbeck, H. Maeng, E. Ogundimu, & K. Perrakis (Eds.), *Developments in statistical modelling* (pp. 1-12). Cham, Switzerland: Springer. [https://doi.org/10.1007/978-3-031-65723-8\\_31](https://doi.org/10.1007/978-3-031-65723-8_31)
- RADINI, R., ALTAROCCA, F., CAPUTI, M., DE FAUSTI, F., FRANCONI, L., ICHIM, D., MANTEGARI, G. (2021, November 30). Come i dati di telefonia mobile possono contribuire alle statistiche sperimentali sulla mobilità [Poster presentation]. Quattordicesima Conferenza Nazionale di Statistica, Roma, Italy.
- RICCIATO, F. (2018). Towards a reference methodological framework for processing MNO data for official statistics. Paper presented at the 15th Global Forum on Tourism Statistics, Cusco, Peru. <https://doi.org/10.5281/zenodo.10396056>
- RICCIATO, F., & COLUCCIA, A. (2023). On the estimation of spatial density from mobile network operator data. *IEEE Transactions on Mobile Computing*, 22(6), 3541-3557. <https://doi.org/10.1109/TMC.2021.3134561>
- RICCIATO, F., DE MEERSMAN, F., WIRTHMANN, A., SEYNAEVE, G., & SKALLOTIS, M. (2018). Processing of Mobile Network Operator data for official statistics: The case for public-private partnership. Paper presented at the DGINS Conference 2018, Bucharest, Romania. <https://doi.org/10.5281/zenodo.10246468>
- RICCIATO, F., LANZIERI, G., WIRTHMANN, A., & SEYNAEVE, G. (2020). Towards a methodological framework for estimating present population density from mobile network operator data. *Pervasive and Mobile Computing*, 68, 101263. <https://doi.org/10.1016/j.pmcj.2020.101263>
- RICCIATO, F., SIIL, T., TALVISTE, R., KUBO, B., & WIRTHMANN, A. (2021). A proof-of-concept solution for the secure private processing of longitudinal Mobile Network Operator data in support of official statistics (Technical report). Eurostat, Unit A5. Retrieved from [https://cyber.ee/uploads/unece2021\\_estat\\_cybernetica\\_v6\\_a03ba175ae.pdf](https://cyber.ee/uploads/unece2021_estat_cybernetica_v6_a03ba175ae.pdf)
- SALUVEER, E., RAUN, J., TIRU, M., ALTIN, L., KROON, J., SNITSARENKO, T., AASA, A., & SILM, S. (2020). Methodological framework for producing national tourism statistics from mobile positioning data. *Annals of Tourism Research*, 81, 102895. <https://doi.org/10.1016/j.annals.2020.102895>
- SCOTTI, F., FLORI, A., SECCHI, P., ARENA, M., & AZZONE, G. (2024). Exploring drivers of overnight stays and same-day visits in the tourism sector. *Scientific Reports*, 14, 9840. <https://doi.org/10.1038/s41598-024-60229-w>
- TARTAGLIA, M., VANNACCI, L., ASTORI, G., FARSI, M., & RADINI, R. (2024). Methodological aspects for passenger mobility analysis using big data (ISBN 978-88-940638-8-2). Ferrovie dello Stato Italiane (FS Research Centre).

- UNITED NATIONS COMMITTEE OF EXPERTS ON BIG DATA AND DATA SCIENCE FOR OFFICIAL STATISTICS. (2021). Mobile phone data. Retrieved from <https://unstats.un.org/bigdata/task-teams/mobile-phone/index.cshtml>
- XIONG, C., HU, S., YANG, M., YOUNES, H., LUO, W., GHADER, S., & ZHANG, L. (2020). Mobile device location data reveal human mobility response to state-level stay-at-home orders during the COVID-19 pandemic in the USA. *Journal of the Royal Society Interface*, 17(173), 20200344. <https://doi.org/10.1098/rsif.2020.0344>
- ZAMENGO, B., DI SORTE, D., CAPPELLARI, D., & ZARAMELLA, A. (2024). I big data per la valorizzazione del turismo montano: esperienze e potenzialità. *Annali del turismo* (Anno XIII). <https://doi.org/10.20373/2283-3102/162>
- ZHENG, Y. (2015). Methodologies for cross-domain data fusion: An overview. *IEEE Transactions on Big Data*, 1(1), 16-34.

**UNDERSTANDING HOTEL COOPERATION IN VENICE:  
A NETWORK ANALYSIS OF KEY DETERMINANTS**

**COMPRENDERE LA COOPERAZIONE TRA HOTEL A VENEZIA:  
UN'ANALISI DI RETE DEI FATTORI CHIAVE**

*Viviana Amati<sup>1</sup>, Anna Moretti<sup>2</sup>, Rebecca Micol Finzi<sup>3</sup>*

**Sommario**

Le destinazioni turistiche sono sistemi di relazioni in cui diversi attori (ad esempio, alberghi e strutture ricettive simili, servizi e gastronomia) offrono attività simili o complementari. Gli attori sono interdipendenti poiché devono cooperare per offrire un prodotto completo ai visitatori: l'esperienza turistica. In questo lavoro si analizzano i dati relativi alle collaborazioni formali e informali tra gli hotel di Venezia. I dati sulle collaborazioni formali sono stati raccolti attraverso le associazioni imprenditoriali locali. Per quanto riguarda le collaborazioni informali, è stato somministrato un questionario utilizzando un disegno basato sul generatore di nomi, chiedendo agli hotel con chi avessero collaborato prima e dopo la pandemia. Sono state anche raccolte informazioni sull'importanza, l'obiettivo e la frequenza delle relazioni. Tutte queste informazioni sono state integrate intervistando gli hotel che non hanno compilato il sondaggio o lo hanno compilato parzialmente. I dati sono stati analizzati per determinare i fattori che spiegano l'importanza e la frequenza della collaborazione tra gli hotel di Venezia, con particolare attenzione alla dotazione di risorse, alla posizione geografica e alle variabili gestionali. Le analisi mostrano che i modelli di collaborazione influenzano la gestione del territorio per uno sviluppo sostenibile.

**Abstract**

*Tourism destinations are networked systems where different actors (e.g., hotels and similar accommodations, services, and gastronomy) offer similar or complementary activities. The actors are interdependent since they must cooperate to deliver a comprehensive product to their visitors: the tourism experience. In this work, we analyze data about formal and informal collaborations among hotels in Venice.*

<sup>1</sup> University of Milano-Bicocca, Department of Statistics and Quantitative Methods, Milan, Italy - e-mail: viviana.amati@unimib.it (corresponding author).

<sup>2</sup> Ca' Foscari University of Venice, Venice School of Management, Venice, Italy - e-mail: anna.moretti@unive.it.

<sup>3</sup> University of Milano-Bicocca, Department of Informatics, System and Communication, Milan, Italy - e-mail: r.finzi1@campus.unimib.it.

*We collected information about formal collaborations through local business associations. For the informal collaborations, we administered a survey using a recall design based on a name generator, asking hotels with whom they collaborated before and after the pandemic. We also collected information on relationships' importance, goal, and frequency. We complement this information by interviewing the hotels that did not or partially fill in the survey. We apply weighted ERGMs to investigate the determinants of the importance and frequency of collaboration among the hotels in Venice, with a focus on resource endowment, geographical location, and management variables. We also describe how patterns of collaboration affect land management for sustainable development.*

## 1. Introduction

Tourism constitutes a key driver of economic development, particularly in Italy, where it accounts for approximately 10.8% of national GDP and supports nearly 13% of total employment (ENIT data). The sector's sustained success and global competitiveness are contingent upon effective governance – defined as the coordination, interaction, and cooperation among various stakeholders within a destination, including entities from the third sector (Hsu *et al.*, 2013). In contemporary tourism systems, achieving both competitiveness and sustainability necessitates collaborative action among stakeholders, rather than isolated or fragmented efforts (Fyall & Garrod, 2020).

From a network-oriented perspective, tourism destinations can be understood as complex, interconnected systems composed of diverse actors – such as hotels, service providers, restaurants, cultural institutions, and other enterprises – that deliver either complementary or substitutable services (Elvekrok *et al.*, 2022). These entities are functionally interdependent, collectively shaping and sustaining the “tourism experience,” which encompasses the full spectrum of a visitor's journey, including accommodation, dining, transportation, entertainment, and cultural engagement.

Collaboration among stakeholders often manifests through mechanisms such as information exchange, mutual assistance, joint training initiatives, and the sharing of resources. Empirical research consistently demonstrates that such cooperative behaviors yield substantial benefits, both at the destination level and for individual organizations (e.g., Elvekrok *et al.*, 2022; Ledesma González *et al.*, 2021; Mwesiumo & Halpern, 2019). Specifically, collaboration contributes to enhanced destination competitiveness (Baggio *et al.*, 2010; Barratt, 2004; Polenske, 2004) and strengthens organizational capabilities by enabling strategic alliances and the pooling of knowledge and resources (Della Corte & Aria, 2016). The sharing of such intangible assets, therefore, constitutes a deliberate and strategic organizational choice (Wilke *et al.*, 2019).

Nonetheless, fostering collaboration within the tourism sector is inherently challenging due to the heterogeneity of stakeholders, which often includes actors from hospitality, transport, food and beverage, and cultural and creative industries. Effective collaboration requires the development of trust, shared understanding, and the alignment of potentially divergent interests (Czernek-Marszałek, 2019; Nunkoo, 2017). Consequently, despite the recognized advantages of collaborative networks, such arrangements remain relatively uncommon and frequently lack long-term stability in many destination contexts (Marasco *et al.*, 2018).

This study focuses on Venice, a world-renowned tourism destination currently facing critical pressures associated with overtourism (Baldin *et al.*, 2024). In this setting, collaboration among local stakeholders is not merely advantageous but essential for advancing sustainable tourism development (Moretti *et al.*, 2024). The present research conceptualizes inter-organizational collaboration in Venice's hospitality sector as a network composed of hotels and the collaborative relationships between them. Employing social network analysis, the study investigates the structural and contextual factors that influence the formation of cooperative ties and identifies the mechanisms that support the emergence and maintenance of collaboration within this urban tourism ecosystem.

## 2. Background

In this study, cooperation among tourism firms is defined in accordance with Czernek's (2013) conceptualization of collaboration as "a form of voluntary joint actions in which autonomous stakeholders engage in an interactive process, using shared rules, norms, and structures, to act and decide on issues related to tourism development in the region" (p. 84). This definition underscores the voluntary and structured nature of collaborative processes, which are essential for effective destination governance.

The emergence and evolution of inter-firm cooperation within the tourism sector have been examined through a variety of theoretical frameworks. These include transaction cost economics (TCE) (Williamson, 1975), the resource-based view (RBV) (Prahalad & Hamel, 1990), resource dependency theory (RDT) (Pfeffer & Salancik, 1978), and theories of social embeddedness (Granovetter, 1973, 1985). Each of these perspectives contributes to understanding the motivations and mechanisms underlying collaborative behavior among firms and the network emergence (Padgett & Powell, 2012). TCE posits that firms pursue cooperative relationships when market transactions incur high costs, making relational governance more efficient. The RBV and RDT, by contrast, emphasize the strategic intent of firms to leverage and complement their internal resources through access to external assets. Meanwhile, the sociological perspective on embeddedness, particularly as articulated by Granovetter (1985), highlights the extent to which

firms are situated within networks of social relationships that both constrain and enable their strategic actions. These networks are not merely contextual but are actively shaped and utilized by firms to achieve organizational goals.

Building on these theoretical foundations, prior research has predominantly concentrated on two key areas: the benefits of collaboration (Elvekrok *et al.*, 2022; Jamal & Getz, 1995; Landorf, 2009; Novelli *et al.*, 2006; Tribe & Liburd, 2016), and the determinants of inter-organizational cooperation (Czernek, 2013; Czernek-Marszałek, 2018, 2020, 2021; López-Molina & Pulido-Fernández, 2024; Moretti *et al.*, 2024). A nuanced understanding of the dynamics underpinning tourism firm cooperation is vital, as collaboration has been shown to enhance organizational competitiveness and contribute to the success and resilience of tourism destinations (Baggio, 2020b; Moretti *et al.*, 2024; Raisi *et al.*, 2020).

This study aligns with the second stream of research, focusing on the determinants of inter-firm collaboration. These determinants are conceptualized as the conditions and factors that influence the initiation, development, and outcomes of cooperative relationships in tourism contexts (Czernek, 2013). They can be broadly classified into four categories: (1) preconditions for collaboration, (2) factors necessary for the successful development of cooperation, (3) individual capabilities critical for managing cooperative relationships, and (4) contextual and structural elements that influence the effectiveness of cooperation (Czernek, 2013). Within these categories, previous research has examined a variety of variables, including the social environment (e.g., embeddedness) (Czernek-Marszałek, 2021), geographical proximity (Owoo & Naudé, 2017), governance structures (Flagestad & Hope, 2001), the presence of intermediary organizations such as trade associations (Moretti *et al.*, 2024), and organizational characteristics such as resource endowments, often proxied by hotel classification levels (Rodríguez-Díaz & Espino-Rodríguez, 2006; Wilke *et al.*, 2019).

Despite these advances, the literature on the determinants of tourism cooperation remains relatively underdeveloped (Beritelli, 2011). Although several frameworks have been proposed (e.g., Czernek, 2013), their application in empirical research – particularly through network-based methodologies – remains limited (Czernek-Marszałek, 2019, 2021). Most existing studies rely on qualitative approaches or utilize basic network statistics to describe structural features, often treating determinants in isolation. However, as determinants of cooperation tend to operate concurrently and interactively, it is essential to adopt methodological approaches that allow for their joint analysis. Furthermore, there has been insufficient attention to endogenous network processes – those in which the existence of one tie influences the likelihood of others forming – which are critical for understanding the mechanisms of network evolution and tie dependency.

To address these gaps, the present study aims to deepen the understanding of cooperation determinants by applying network models to examine the question: *What factors influence collaboration among hotels in Venice?* Employing a network-based analytical framework grounded in statistical modeling, this research investigates the structural interdependencies inherent in collaborative ties among hospitality firms. This approach not only accommodates the relational nature of network data but also enables the simultaneous evaluation of individual and structural determinants of collaboration. In doing so, it offers a more comprehensive perspective on the drivers of inter-organizational cooperation within urban tourism networks.

### 3. Data

We collected data about all hotels, B&Bs, hostels, and apartments in Mestre and Venice, using the roster provided on the Open Data website of the Veneto region (<https://www.regione.veneto.it/web/informatica-e-e-government/open-data>). This website offers a comprehensive data set including all the hospitality firms including hotels, B&Bs, hostels, and apartments, along with details about their facilities (such as private parking, restaurants, and swimming pools), quality and standards (expressed through star ratings and the “Leoni” system), and their geographic locations (specified by the addresses).

From this dataset, we computed pairwise facility-based similarities between hotels using the Jaccard similarity index. Geographic distances were calculated using the geodesic (great-circle) distance in kilometers, after deriving latitude and longitude coordinates from each hotel’s address.

Data on inter-firm collaboration were obtained through an online survey administered in 2022 by the Venice School of Management at Ca’ Foscari University of Venice. All registered hospitality firms in Mestre and Venice were contacted via email to participate. The survey adopted a name generator and recall method, asking respondents to identify up to five collaborators across several domains, including other accommodation providers, food service businesses, cultural institutions, and miscellaneous partners. Respondents were also asked to evaluate the significance of each collaboration on a scale from 1 (not important at all) to 5 (very important) and to specify the primary motivations behind the collaboration. For the purposes of this study, only collaborations between hotels were analyzed.

Additionally, data on formal affiliations – defined as memberships in local or national business associations – were collected by asking respondents to list any regional or national hotelier organizations they belonged to. To enhance survey coverage, telephone interviews were conducted with hotels that had not begun the questionnaire or had provided only partial responses. Outreach followed a priority order based on star rating, starting with four- and five-star establishments and proceeding to three-star ho-

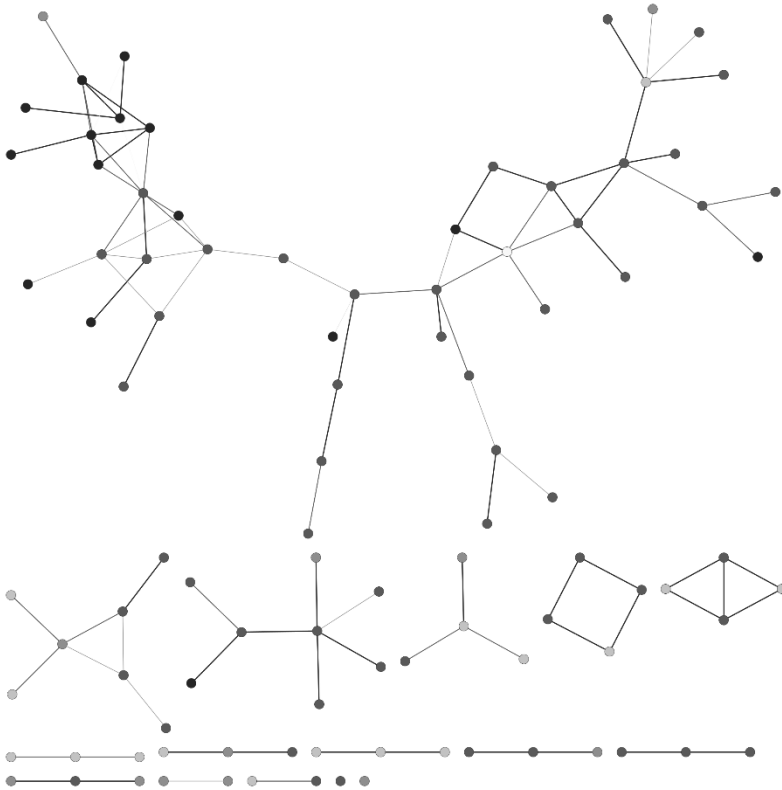
tels; however, resource constraints prevented the extension of this follow-up phase to one- and two-star properties.

Figure 1 shows an excerpt of the network of collaboration ties among the 122 hotels out of 455 (response rate 26.81%) that were part of the final sample. In this graph, each node represents a hotel, while each link signifies a collaboration tie. The intensity of the node colors corresponds to hotel quality: 72.95% (89 hotels, out of 159, response rate 55.97%) have four or five stars, 21.31% (26 hotels, out of 196, response rate 13.27%) have three stars, and 5.74% (7 hotels, out of 100, response rate 7.00%) have one or two stars. The distance between hotels ranges from 0.4 to 24.9 kilometers, with a median and average distance of 2.1 and 6.5 kilometers, respectively. The similarity in terms of offered facilities (not shown in Figure 1) varies from 0.15 to 1, with an average similarity of 0.5.

The network comprises 177 reported collaboration ties, with an average importance rating of 3.8 (on a 1–5 scale). The distribution of importance scores is notably skewed. The primary motivations for collaboration include innovation and the development of new products and services (41.9%, 74) and the management of overbooking (33.1%, 59). Less frequently cited motivations include joint marketing initiatives (19.8%, 35) and staff training (5.2%, 9).

The collaboration network is characterized by low density (0.024) – indicating that only 2.4% of all possible ties are realized – and includes a few disconnected components. The degree distribution – defined as the number of collaborators per hotel – ranges from 1 to 8, with a median degree of 4 and a mean degree of 2.7, reflecting a right-skewed distribution.

*Figure 1. Collaboration networks among hotels in Venice. The node intensity color is proportional to the hotel quality, while the edge intensity color and width to the importance of the collaboration. For confidentiality reasons, only a subgraph of the full network is displayed and node identifiers have been omitted to prevent the identification of individual hotels*



Source: Open Data website of the Veneto region and survey data

#### 4. Methods

To investigate the determinants of collaboration among hotels in Venice – specifically focusing on resource endowment, geographic location, management characteristics, and formal relationships – we employ valued exponential random graph models (ERGMs).

ERGMs (Wasserman & Pattison, 1996; Robins *et al.*, 2010; Lusher *et al.*, 2013) are statistical models designed for analyzing a single observation of a network. While initially developed for binary networks (where ties are either present or absent), extensions to ERGMs family allow for the modeling of valued networks, where ties are assigned a weight representing the intensity of the relationship (Krivitsky, 2012).

Let  $\mathbf{x}$  represent the observed network as an adjacency matrix, where the rows and the columns are the hotels, and the element  $x_{ij}$  reports the intensity of the relationships between hotel  $i$  and  $j$ . We denote the monadic and dyadic covariates by  $v$  and  $w$ , encoding the characteristics of the hotels (e.g., stars or number of offered facilities) and pairs of hotels (e.g., geographical distance and facility similarity), respectively.







The valued ERGM assumes that the observed network  $x$  arises from a set of underlying social mechanisms driving tie formation. In our context, these mechanisms include: the baseline tendency of establishing collaborations and their importance; *transitivity*, the tendency of hotels to collaborate with the collaborators of their collaborators, or in other words, the tendency for two hotels to be connected if they share a common collaborator; *homophily*, the tendency of hotels to collaborate with hotels that have similar characteristics; the *activity* the tendency of hotels having a certain attribute to collaborate with more partners than hotels having a different characteristic; *proximity*, the tendency of hotels to collaborate with nearby hotels. Those mechanisms act simultaneously and ERGMs allow to parse out the effect of each mechanism.

The valued ERGM takes the form:

$$P(X = x) = h(x) \frac{\exp\{\sum_k \vartheta_k s_k(x, v, w)\}}{k(\vartheta)} \quad (1)$$

with  $h(x)$  the reference distribution,  $s_k(x, v, w)$  the statistics playing the role of explanatory variables  $\vartheta_k$ , parameters expressing the importance of the statistics in explaining the structure of the observed network, and  $k(\vartheta)$  is a normalizing constant. The statistics  $s_k(x, v, w)$  are counts of local configurations of ties representing mechanisms that have generated the observed network. Their choice is theory-driven, i.e. guided by substantive theories and hypotheses concerning the analyzed phenomenon.

Table 1. Mechanisms, local configurations and statistics for valued ERGMs

Mechanism	Local configurations	Statistic $s_k(\mathbf{x}, v, w)$
Ties		number of collaboration ties
Intensity		sum of the importance values
Transitivity		sum of transitive triads
Activity		sum of the values of ties weighted by the hotel attribute
Homophily		sum of the values of ties weighted by the hotels similarity
Proximity		sum of the values of ties weighted by geographical proximity

Intuitively, the model in Equation (1) expresses the global structure of the network, denoted as  $P(X=\mathbf{x})$ , as a function of local configurations  $s_k(\mathbf{x}, v, w)$  which are empirical representations of underlying tie formation mechanisms. The influence of each configuration is captured by the parameter  $s_k(\mathbf{x}, v, w)$ , which must be estimated from the observed data.

Table 1 provides definitions and computational procedures for each network statistic, along with visual representations (pictograms) of the corresponding local configurations. Statistics associated with *ties*, *intensity*, and *transitivity* capture endogenous mechanisms, reflecting the dependency of a given tie on the presence of other ties within the network. These are thus referred to as *endogenous statistics*. In contrast, the statistics representing *activity*, *homophily*, and *proximity* account for the influence of node-level attributes – such as hotel characteristics – on tie formation and are accordingly classified as *exogenous statistics*.

## 5. Results

Table 2 presents the estimated parameters of the specified Exponential Random Graph Models (ERGMs). With the exception of the activity-related variables, all parameters are statistically significant at the 5% level.

Table 2. Valued ERGMs for the importance of collaboration among hotels. Parameters estimates (Est.), standard errors (s.e.), and significance (Sign.)

Statistic	Est.	s.e.	Sign.
<i>Endogenous effects</i>			
Ties	-6.16	0.3	***
Intensity	-0.61	0.31	*
Transitivity	0.52	0.09	***
<i>Exogenous effects</i>			
Stars activity	0.11	0.18	
Stars similarity (absolute difference)	-0.19	0.09	*
Facility activity	0.14	0.16	
Facility similarity	0.16	0.07	*
Same business association	0.49	0.13	***
Geographical distance (logarithm)	-0.07	0.03	*
***p < 0.001; **p < 0.01; *p < 0.05;			

Source: Own analysis

We begin by discussing the endogenous structural effects. The negative coefficients associated with the *ties* and *intensity* statistics indicate a sparse network structure, characterized by a limited number of collaboration links. This result is consistent with the observed low network density, where only 3% of all potential ties are realized. Furthermore, the positive and significant coefficient for *transitivity* suggests a propensity for triadic closure; that is, hotels are more likely to establish collaboration ties when they share a common partner, reflecting a tendency toward clustering within the network.

Turning to the exogenous covariate effects, the coefficient for *stars activity* – which captures the effect of hotel quality (measured by star rating) on the likelihood of forming ties – is positive but not statistically significant. This implies that collaboration propensity does not differ significantly between higher- and lower-quality hotels. Conversely, the negative and significant coefficient for *stars similarity*, operationalized as the absolute difference in star ratings between hotel pairs, indicates a homophilous pattern: hotels are more inclined to collaborate with others of similar quality.

Regarding service offerings, the coefficient for the *activity effect* based on the number of facilities provided is likewise positive but not statistically significant, suggesting that the volume of services offered does not significantly influence the likelihood of collaboration. However, the positive and significant coefficient for *facility similarity* points

to a tendency for hotels to form ties with others offering similar services, indicating that alignment in service provision fosters collaboration.

Finally, the results underscore the importance of institutional and spatial proximity. The positive and significant coefficient for *shared business association membership* highlights the facilitating role of formal organizational affiliations in promoting inter-hotel collaboration. In contrast, the negative and significant coefficient for *geographical distance* indicates that spatial proximity increases the likelihood of tie formation, underscoring the role of geographic closeness in shaping collaborative behavior.

## 6. Conclusions and discussions

This study investigates the determinants of collaboration among hospitality firms in Venice, with particular attention to the significance attributed to these relationships. The results of the model highlight the central role of transitivity in shaping collaborative patterns within the hotel sector. Specifically, the positive and statistically significant parameter associated with transitivity suggests that hotels are inclined to form tightly knit clusters, whereby collaboration tends to occur within small, interconnected groups. This finding aligns with broader network theory, which identifies triadic closure as a fundamental mechanism of social capital accumulation (Granovetter, 1973): once two hotels share a common collaborator, the likelihood of a direct tie forming between them increases substantially. From a governance perspective, this has a practical implication: even targeted interventions that connect a small number of otherwise isolated hotels to existing clusters could trigger a cascade of new collaborative ties, amplifying the overall density and resilience of the destination network.

Another key driver identified is *homophily* – the tendency for entities to associate with similar others. The findings indicate that hotels exhibit a preference for collaborating with establishments of comparable quality and service offerings. This homophilous behavior may be partially explained by operational practices such as overbooking management, wherein hotels redirect guests to alternative accommodations that match their own quality standards. Furthermore, similarity in attributes such as management style (e.g., professionally managed versus family-run operations) may foster a shared organizational culture or “language,” thereby facilitating trust, communication, and cooperation (Nunkoo, 2017). While homophily facilitates collaboration within quality groups, it simultaneously risks producing a segmented network structure in which the hotel categories determine the formation of dense and self-reinforcing clusters.

The analysis also underscores the significance of business association membership as a determinant of collaboration. Such associations often provide institutional platforms that promote interaction and trust-building among members, thereby enhancing the likelihood of forming collaborative ties. This finding supports earlier research on the

role of intermediary institutions in fostering cooperation within the tourism and hospitality sectors (Moretti *et al.*, 2024). This result suggests that formal organizational affiliation operates as a structural shortcut to trust: rather than relying solely on the slow accumulation of relational experience between individual hotels, membership in a shared association provides a pre-existing institutional context that lowers the perceived risk of cooperation.

Geographical distance likewise emerges as a critical factor in the formation of collaborative relationships. The unique urban and territorial morphology of Venice – characterized by its compact historic center and fragmented geography across more than a hundred islands – intensifies the relevance of spatial proximity. In this context, short distances take on greater practical importance than in other urban settings. As a result, collaborative practices such as overbooking accommodations are often restricted to hotels located in close physical proximity, reinforcing the salience of geographic closeness in the formation of collaborative ties. The concentration of collaborative ties within geographically proximate clusters suggests that the benefits of cooperation (including guest redistribution, shared resources, and joint problem-solving) may not diffuse evenly across the destination. In particular, hotels located in peripheral areas or on less-connected islands may find themselves structurally disadvantaged in terms of access to collaborative partners.

This study represents a preliminary attempt to analyze the collaborative dynamics among hotels in Venice and may be regarded as a pilot investigation to inform future, more comprehensive research. However, certain limitations must be acknowledged. First, the overall response rate was 26.81% (122 out of 455 hotels). Response rates varied considerably across star categories: 55.97% among four- and five-star hotels (89 out of 159), 13.27% among three-star hotels (26 out of 196), and 7.00% among one- and two-star hotels (7 out of 100). These differences reflect both the priority-based follow-up design and the generally lower engagement of smaller, lower-category establishments with survey-based research. Consequently, findings should be interpreted with caution with respect to the generalizability to one- and two-star properties, which remain underrepresented in the final sample. Second, the presence of missing data, particularly non-responses that result in unobserved ties, may have altered the observed network structure and potentially influenced the estimation and interpretation of model parameters. As such, the results should be interpreted as exploratory and descriptive in nature and cannot be generalized to the broader population of hospitality firms in Venice.

Additionally, the dataset relies on a single cross-sectional survey conducted in 2022, which limits our ability to assess changes in network dynamics over time. The post-pandemic recovery period would have been an important opportunity to observe potential structural shifts in inter-hotel collaboration. However, the lack of longitudinal data pre-

vents any comparison between pre- and post-pandemic conditions, as well as a temporal validation of our findings. This decision was influenced by practical and resource constraints related to primary data collection within the specific geographic area.

Finally, the analysis focuses exclusively on hotels, thereby excluding other stakeholders, such as food and beverage businesses, transport providers, and cultural institutions, that actively participate in the Venetian tourism ecosystem. This boundary condition was primarily imposed by data availability constraints: comprehensive and reliable contact lists for non-hotel tourism operators in the study area were not accessible at the time of data collection, and preliminary outreach attempts to broaden the sample yielded insufficient responses to warrant inclusion. Future research could address this gap by developing dedicated sampling frames for other stakeholder categories, potentially through collaboration with local tourism authorities or business associations, to provide a more complete picture of the relational structure underpinning destination competitiveness.

Despite these limitations, this study offers valuable insights into the mechanisms underlying inter-organizational collaboration in urban tourism networks. It provides a conceptual and methodological framework that may be applied to other geographic and sectoral contexts to further explore the structural and contextual drivers of collaboration among hospitality firms.

### **Acknowledgements**

We thank Federica Donai, Romina Vian, Alessia Vidali for their work on data collection during their internship at the University Ca' Foscari University of Venice.

### **Bibliografia**

- BAGGIO, R. (2020b). Tourism destinations: A universality conjecture based on network science. *Annals of Tourism Research*, 82.
- BAGGIO, R., SCOTT, N., COOPER, C. (2010). Improving tourism destination governance: a complexity science approach. *Tourism Review*, 65(4), 51-60.
- BALDIN, A., BERTOCCHI, D., CAMATTI, N. (2024). Crowding-in and (temporary) crowding-out in Venice. The effect of cultural events on residents. *Tourism Economics*, <https://doi.org/10.1177/13548166241233773>.
- BARRATT, M. (2004). Understanding the meaning of collaboration in the supply chain. *Supply Chain Management: an international journal*, 9(1), 30-42.
- BERITELLI, P. (2011). Cooperation among prominent actors in a tourist destination. *Annals of Tourism Research*, 38(2), 607-629.

- CZERNEK-MARSZAŁEK, K. (2021). The sources and components of social embeddedness as determinants of business cooperation in a tourist destination. *Journal of Destination Marketing & Management*, 19, 100534.
- CZERNEK-MARSZAŁEK, K. (2020). The overembeddedness impact on tourism cooperation. *Annals of Tourism Research*, 81, 102852.
- CZERNEK-MARSZAŁEK, K. (2019). Applying mixed methods in social network research – The case of cooperation in a Polish tourist destination. *Journal of Destination Marketing & Management*, 11, 40-52.
- CZERNEK-MARSZAŁEK, K. (2018). Cooperation evaluation with the use of network analysis. *Annals of Tourism Research*, 72, 126-139.
- CZERNEK, K. (2013). Determinants of cooperation in a tourist region. *Annals of Tourism Research*, 40, 83-104.
- DELLA CORTE, V., ARIA, M. (2016). Coopetition and sustainable competitive advantage. The case of tourist destinations. *Tourism Management*, 54, 524-540.
- ELVEKROK, I., VEFLÉN, N., SCHOLDERER, J., SØRENSEN, B.T. (2022). Effects of network relations on destination development and business results. *Tourism Management*, 88, 104402.
- FLAGESTAD, A., HOPE, C.A. (2001). Strategic success in winter sports destinations: a sustainable value creation perspective. *Tourism Management*, 22(5), 445-461.
- FYALL, A., GARROD, B. (2020). Destination management: a perspective article. *Tourism Review*, 75(1), 165-169.
- GRANOVETTER, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3), 481-510.
- GRANOVETTER, M. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6), 1360-1380.
- HSU, T.W., INBAKARAN, R., GEORGE, B. (2013). Understanding the concept of governance in tourism policy-making. *International Journal of Economics and Business Research*, 5(3), 319-336.
- JAMAL, T.B., GETZ, D. (1995). Collaboration theory and community tourism planning. *Annals of Tourism Research*, 22(1), 186-204.
- KRIVITSKY, P.N. (2012). Exponential-family random graph models for valued networks. *Electronic journal of statistics*, 6, 1100.
- LANDORF, C. (2009). Managing for sustainable tourism: a review of six cultural World Heritage Sites. *Journal of Sustainable Tourism*, 17(1), 53-70.
- LEDESMA GONZÁLEZ, O., MERINERO-RODRÍGUEZ, R., PULIDO-FERNÁNDEZ, J.I. (2021). Tourist destination development and social network analysis: What does degree centrality contribute?. *International Journal of Tourism Research*, 23(4), 652-666.
- LÓPEZ-MOLINA, D., PULIDO-FERNÁNDEZ, J.I. (2024). Analysis of the interaction and collaboration

- of the actors of a tourist destination from a network perspective. *Current issues in tourism*, 27(10), 1572-1589.
- LUSHER, D., KOSKINEN, J., ROBINS, G. (2013). *Exponential random graph models for social networks: Theory, methods, and applications*. Cambridge University Press.
- MARASCO, A., DE MARTINO, M., MAGNOTTI, F., MORVILLO, A. (2018). Collaborative innovation in tourism and hospitality: a systematic review of the literature. *International Journal of Contemporary Hospitality Management*.
- MORETTI, A., MARTINI BARZOLAI, M., CUTUGNO, M. (2024). Ambidexterity of hospitality trade associations: a social capital perspective for destination coordination and cooperation dynamics. *Current issues in tourism*, 1-22.
- MWESIUMO, D., HALPERN, N. (2019). A review of empirical research on interorganizational relations in tourism. *Current Issues in Tourism*, 22(4), 428-455.
- NOVELLI, M., SCHMITZ, B., & SPENCER, T. (2006). Networks, clusters and innovation in tourism: A UK experience. *Tourism Management*, 27(6), 1141-1152.
- NUNKOO, R. (2017). Governance and sustainable tourism: What is the role of trust, power and social capital? *Journal of Destination Marketing & Management*, 6(4), 277-285.
- OWOO, N.S., NAUDÉ, W. (2017). Spatial proximity and firm performance: evidence from non-farm rural enterprises in Ethiopia and Nigeria. *Regional Studies*, 51(5), 688-700.
- PADGETT, J.F., POWELL, W.W. (2012). The Problem of Emergence. In PADGETT, J.F., POWELL, W.W. (Eds.), *The emergence of organizations and markets* (pp. 1-29). Princeton: Princeton University Press.
- PFEFFER, J., SALANCIK, G. R. (1978). The external control of organizations: A resource dependence approach. *NY: Harper and Row Publishers*.
- PRAHALAD, C., HAMEL, G. (1990). The Core Competence of the Corporation. *Harvard Business Review*, 68(3), 295-336.
- POLENSKE, K. R. (2004). Competition, collaboration and cooperation: an uneasy triangle in networks of firms and regions. *Regional Studies*, 38(9), 1029-1043.
- RAISI, H., BAGGIO, R., BARRATT-PUGH, L., WILLSON, G. (2020). A network perspective of knowledge transfer in tourism. *Annals of Tourism Research*, 80, 102817.
- ROBINS, G., PATTISON, P., KALISH, Y., LUSHER, D. (2007). An introduction to exponential random graph ( $p^*$ ) models for social networks. *Social networks*, 29(2), 173-191.
- RODRÍGUEZ-DÍAZ, M., ESPINO-RODRÍGUEZ, T.F. (2006). Developing relational capabilities in hotels. *International Journal of Contemporary Hospitality Management*, 18(1), 25-40.
- TRIBE, J., & LIBURD, J.J. (2016). The tourism knowledge system. *Annals of Tourism Research*, 57, 44-61.
- WASSERMAN, S., PATTISON, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and  $p^*$ . *Psychometrika*, 61(3), 401-425.

- WILKE, E., COSTA, B., FREIRE, O., FERREIRA, M. (2019). Interorganizational cooperation in tourist destination: Building performance in the hotel industry. *Tourism Management*, 72, 340-351.
- WILLIAMSON, O.E. (1975). Markets and hierarchies: analysis and antitrust implications: a study in the economics of internal organization. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*. Free Press.

## TWIN TRANSITION, SPECIALIZZAZIONI INTELLIGENTI E NUOVE GEOGRAFIE DELL'INNOVAZIONE

### TWIN TRANSITION, SMART SPECIALISATIONS AND NEW GEOGRAPHIES OF INNOVATION

*Sergio Salamone<sup>1</sup>, Francesco Giovanni Truglia<sup>2</sup>*

#### **Sommario**

In questo articolo viene esaminata la relazione tra aree di specializzazione intelligente e il loro contributo allo sviluppo territoriale, anche in contesti meno sviluppati e periferici, soprattutto in termini di innovazione digitale e transizione verde. In quali territori questa relazione è diffusa o assente, pur in presenza di buoni livelli di digitalizzazione e sostenibilità?

In primo luogo, viene fornita una breve panoramica del contesto di riferimento, per poi illustrare le fonti di riferimento dei dati e le metodologie adottate. I risultati mostrano innanzitutto la distribuzione dei livelli di specializzazione intelligente nei territori Nuts3 (nomenclatura delle unità territoriali statistiche) e l'associazione tra digitalizzazione e sostenibilità ambientale e sociale. Successivamente, viene presentata un'analisi della correlazione spaziale tra i domini della doppia transizione (indice digit e green) e i legami con la specializzazione intelligente (punteggio dell'indice S3). Sono stati ricercati cluster di territori che presentano similarità o effetti di contaminazione sui tre domini considerati. Infine, è stata proposta un'analisi sperimentale a livello territoriale sulle dodici aree di specializzazione italiane.

#### **Abstract**

*This paper aims to examine the relationship between smart specialisation areas and their contribution to territorial development, including in less developed and peripheral regions, with particular emphasis on digital innovation and the green transition. In which territories is this relationship well established, and where is it weak or absent, even in the presence of relatively high levels of digitalization and sustainability?*

*First, a brief overview of the reference context is provided, followed by a description of the data sources and the methodologies adopted. The results initially illustrate*

<sup>1</sup> Istat - Istituto Nazionale di Statistica, Roma, Italia - e-mail: sesalamo@istat.it (corresponding author)

<sup>2</sup> Istat - Istituto Nazionale di Statistica, Roma, Italia - e-mail: truglia@istat.it

*the distribution of smart specialization levels across Nuts3 (Nomenclature of Territorial Units for Statistics) territories and examine the relationship between digitalization and environmental and social sustainability. Subsequently, the spatial correlation between the domains of the twin transition-digital and green, as measured by the respective indices-and their links with smart specialization (S3 index score) is analyzed. We then identify clusters of territories exhibiting similarities or spillover effects across the three domains considered. Finally, an experimental territorial-level analysis focusing on the twelve Italian smart specialization areas is presented.*

## 1. Introduction

The Smart Specialisation Strategy (S3) has been confirmed and strengthened within the 2021–2027 programming period of EU cohesion policy, with the aim of fostering a new European model of smart, sustainable, and inclusive growth. This model is intended to drive the green, digital, and resilient transitions of the economy and society, ultimately enhancing citizens' well-being.

With reference to the twelve smart specialization areas identified in Italy, which encompass nearly all major economic activities, a new enterprise classification was developed. This classification is based on a combination of data from the enterprise census survey and integrated statistical registers, all drawn from Istat (the Italian National Institute of Statistics). The statistical indicators constructed by specialization area indicate stronger performance among smart-specialized enterprises, particularly in terms of exports, value added, investment in research and development, levels of digitalization, and human capital. Moreover, as expected, smart specialization areas are more strongly represented in the north-east and north-west of the country than in the southern regions, and they predominantly involve medium-sized and large enterprises rather than small firms.

To support the adoption of advanced technologies, particularly those associated with the digital transition, and to ensure that enterprises fully benefit from investments in research, development, and innovation (r&d&i), it is necessary to design a set of complementary measures alongside direct investment incentives. These measures should aim to enhance the availability of skills and other intangible assets at the firm level, such as managerial capabilities and the development of inter-firm networks.

The measurement tools proposed in this work also encompass these dimensions (e.g. the human capital indicator and the indicator of inter-firm relationships), which can be analyzed using the same methodological framework applied to the twin transition indicators. This approach enables the identification of territorial patterns as well as specific spillover effects.

This work aims to identify new geographies of innovation by analysing the spatial relationships generated by the activities of smart-specialization firms. These relationships embed latent dimensions of innovation, which are identified and measured through purpose-built indices, such as those capturing digitalization and sustainability. Territorial development trajectories, shaped by spillovers between neighboring provinces, are measured using the smart specialization (S3) index. This composite index also integrates economic performance indicators to capture the wealth generated within each territory, particularly labor productivity, and provides a robust representation of the concentration and impact of smart-specialization firms across territories.

The originality of the work lies in the application of spatial correlation techniques, using indices specifically constructed at the firm level to measure and monitor smart specialization across territories, and to detect productive and innovative relationships in contiguous areas. The indices used provide new insights into territorial relationships compared to standard innovation measures (e.g., patents, publications), as they capture qualitative aspects and firm strategies, for instance, variables related to firm-university collaborations, social sustainability, training, and the level of tertiary employment within firms.

Since the S3 policy is designed around the idea of development based on territorial proximity and related variety as drivers of innovation, this work identifies territories with cognitively similar sectors that foster knowledge and innovation spillovers. Conversely, macro-areas emerge where these forms of innovation are absent, and no spillover effects are created, perhaps due to a distorted implementation of the policy, namely through a non place-based but copy-paste policy approach.

The paper is structured as follows: first, a brief overview of the reference context is provided, followed by a description of the data sources and the methodological approach. The results initially illustrate the distribution of smart specialization levels across Nuts3 territories and explore the relationship between digitalization and environmental and social sustainability. Next, the spatial correlations between the domains of the twin transition-digital and green-and their connections with smart specialization (S3 index) are analyzed. Territorial clusters exhibiting similarities or spillover effects across the three domains are then identified. Finally, an experimental analysis at the territorial level is conducted, focusing on the twelve Italian smart specialization areas.

The presented approach and findings have implications not only for research, development, and innovation policies but also for broader competitiveness policies, including the Smart Specialization Strategy (S3). The analysis identifies specific provinces with high potential development trajectories, due to their proximity to and interaction with more advanced economies, suggesting priority targets for investment in r&d and innovation to foster balanced regional innovation. Conversely, other territories, parti-

cularly in Southern Italy, exhibit limited prospects, characterized by low relatedness, weak external networks, and policies that are often imitative rather than rooted in endogenous potential. This framework supports a move away from “one-size-fits-all” policies toward a differentiated, territorially-sensitive innovation policy approach. Its experimental application, designed to map territorial potential across the twelve areas of specialization and their interrelations, can inform monitoring and evaluation processes, as well as guide the reprogramming of regional smart specialization strategies, thereby providing actionable support for S3 policymaking.

## 2. Reference context

Conceived within the reformed cohesion policy of the European Commission, smart specialization is a place-based approach that identifies strategic areas for intervention based on an analysis of both the strengths and potential of the economy.

The smart specialization approach integrates industrial, educational and innovation policies, encouraging countries or regions to select a limited number of priority areas for knowledge-based investments, focusing on their strengths and comparative advantages.

The connection between smart specialization and the identification of technological domains within territories forms the foundation of the European smart specialization strategy (Foray 2013).

The evolution of innovation geography theory is closely linked to the rise of Smart Specialisation Strategy (S3) policy, which consolidates the place-based approach (Barca, 2009). In particular, to select investment priorities in specific areas: regions must set priorities in specific research and innovation areas where they have a competitive advantage or significant potential, rather than spreading resources thinly across all sectors (Boschma, 2005).

However, territorial proximity alone is not sufficient. Territories diversify more easily toward technologies that are related (related variety) to those already present (Boschma, R. & Frenken, K., 2011). The evolution of the concept of relatedness, applicable at multiple levels, not only technological but also industrial and occupational (Boschma, 2017), reinforces the need for place-based policies as a bottom-up approach to identify areas of territorial competitive advantage based on local capabilities (Balland, Boschma, Crespo & Rigby, 2019).

Scientific research on territorial innovation also includes how networks and relationships form within territories, influencing resilience even in times of crisis, and in particular the importance of the variety of relationships built with universities and research centers, as well as with other firms, which foster knowledge diffusion and cooperation (Ascani, Balland & Morrison, 2020).

The debate on the geography of innovation thus broadens and supports the foundations of the S3 policy, particularly the place-based approach but also attention to well-being, sustainability, and a less imitative view of smart specialization. Innovation is thus discussed as a plural and inclusive territorial phenomenon, attentive to well-being and balanced territorial development (Binz & Castaldi, 2024).

In recent years, the link between smart specialization and environmental sustainability has strengthened, leading to the evolution of S3 policy into a more defined Smart Specialization and Sustainable Strategy (S4) (Miedzinski et al., 2021). Enterprises are increasingly expected to pair technological transformation with a shift towards more sustainable practices. At the institutional level, various initiatives have been launched to align place-based innovation with sustainable development goals.

The aim is to develop a conceptual framework and a tool to monitor smart specialization strategies in relation to the Sustainable Development Goals (SDGs). One key objective is to explore the connection between S3 and sustainability by identifying appropriate monitoring tools: how can progress in the core components of S3 strategies be evaluated with respect to sustainability objectives?

In Italy, twelve priority areas of specialization have been defined: 1) aerospace; 2) agrifood; 3) blue Growth; 4) green chemistry; 5) design, creativity and made in Italy; 6) energy; 7) smart factory; 8) sustainable mobility; 9) health; 10) smart, secure and inclusive communities; 11) technologies for living environments; 12) technologies for cultural heritage.

Each Italian region selects one or more S3 areas and develops a regional smart specialization strategy, which serves as a prerequisite for accessing cohesion funds for R&D and innovation. To support the monitoring of these regional strategies, Istat has introduced a new classification of smart specialization enterprises and developed indicators to track the performance of the identified areas.<sup>3</sup>

### 3. Methodology

#### 3.1 S3 classification and composit indicators

S3 has a multidimensional nature, so we defined a conceptual model that provides a framework for the theoretical definition of S3.

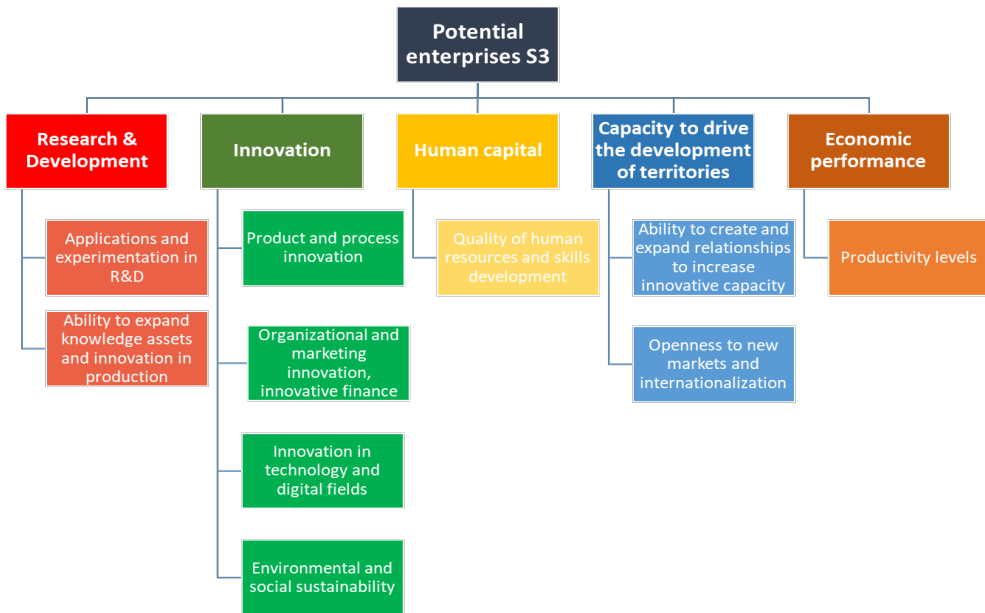
The framework is structured around five key dimensions that represent the general concept of an S3 enterprise: “research and development”, “innovation”, “human capi-

<sup>3</sup> For an overview of the statistical outputs, i.e., methodology used to define S3 classification, composite indicators and tables, visit the following links: <https://www.istat.it/it/statistiche-politiche-sviluppo/specializzazione-intelligente> [https://www.agenziacoesione.gov.it/wp-content/uploads/2022/03/Guida-alla-lettura-degli-indicatori-S3\\_notametodologica-4.pdf](https://www.agenziacoesione.gov.it/wp-content/uploads/2022/03/Guida-alla-lettura-degli-indicatori-S3_notametodologica-4.pdf)

tal”, “capacity to foster local development” and “economic performance” (Figure 1).

Building on this theoretical framework, we operationalized these concepts by linking elementary indicators – constructed from pre-selected elementary variables – to their corresponding sub-factors. This framework serves as the foundation for identifying potential S3 enterprises.

Figure 1. Conceptual framework for the S3 theoretical definition



Source: our elaborations based on data from enterprises census and statistical business registers, 2022, Istat

The S3 composite index was calculated as a synthesis of the five S3 dimensions using the Mazziotta–Pareto index (Mazziotta & Pareto, 2016), a non-compensatory composite index that aggregates sub-dimensions into overall dimensions. This approach ensures non-substitutability of indicators and independence from a so-called “ideal unit,” which is subjective and may vary over time. Sub-dimensions, including the digital and green indices, were calculated using the Wrocław taxonomic method, based on the distance from an ideal unit representing the best performance across all considered indicators. Elementary indicators were derived from the 2022 Istat business census, which included a dedicated S3 section in the survey questionnaire. The scope of the S3 areas was also determined using a census question in which enterprises were asked to self-identify within one or more of the twelve specialization areas (Istat, 2022).

### 3.2 *Univariate and bivariate global and local spatial autocorrelation indices*

The increasing interconnection between digitalization and ecological (green) transition processes, as well as the interactions between smart specialization activities, represents a complex phenomenon with significant social, economic, and cultural implications. In this work, we chose to interpret this dual transition and the relationships between specializations from a territorial perspective, aiming to identify and describe key points of connection between digitalization and sustainable, intelligent production processes.

From this perspective, the territory is not merely a geographic “container” where these connections occur, but also a space of integration between smart specialization processes and multiple local dimensions, including infrastructure endowment, technological capabilities, human capital, collaborative networks, and public policies.

From a territorial perspective, digital and green transition policies can be examined through two complementary lenses. The first concerns territorial embedding, that is, the areas where these practices are already established or in the process of being consolidated. The second focuses on the territorial diffusion of such practices and on the spatial mechanisms that either facilitate or hinder their spread.

From a statistical and methodological perspective, it is therefore necessary to adopt analytical procedures capable of capturing not only the characteristics of individual territories (provinces) but also the interrelations and mutual influences among neighboring areas. To this end, we adopted the approach proposed by Anselin, known as Exploratory Spatial Data Analysis (ESDA), which combines statistical techniques with cartographic visualization to identify patterns, clusters, spatial dependencies, and discontinuities in territorial data.

Specifically, the tools used are the univariate (I) and bivariate (Ib) Moran’s autocorrelation indices, both computed in their global and local versions (Local Indicators of Spatial Association – LISA).

The concept of spatial autocorrelation (or spatial association) stems from the observation that the values of a variable are not independently distributed across a territory; rather, they tend to cluster in specific areas (Demarinis *et al.*, 2011).

Moran’s I index is one of the most widely used measures for assessing the spatial autocorrelation of a variable. In this study, the univariate Moran’s I index is employed to determine whether the territorial configuration, at the Nuts3 level, of smart specialization, digitalization, and green transition processes exhibits similar patterns in neighboring provinces. When this is the case, it indicates an aggregative spatial process and the presence of spatial spillovers, leading to the formation of specialized, digitalized, and/or green territories.

The bivariate version of Moran’s statistic is used to detect spatial interactions between the two types of transition at the level of smart specialization enterprises. Specific-

ly, it evaluates whether the levels of specialization and digital and/or green transition observed in a given province are associated with the levels observed in geographically contiguous provinces. The decomposition of the indices at the local level allows the identification of provinces that make a significant contribution to the aggregation process, thereby forming the core territories of smart specialization.

Overall, spatial analysis based on Moran's indices represents an essential methodological approach for studying the complex relationship between digitalization and ecological transition, highlighting the key role of the territory as a hub for the diffusion and integration of innovative processes. This perspective is crucial for the design of targeted territorial development strategies and the promotion of effective public policies that can leverage spatial dynamics and inter-area spillovers.

The territorial configuration of the variables of interest is analyzed at the Nuts3 level using Moran's spatial autocorrelation index in both its global and local forms (local indicators of spatial association).

### 3.2.1 Global univariate Moran's I

The univariate Moran's I index measures the presence of spatial autocorrelation for a single statistical variable. It indicates whether the territorial configuration of that variable, at the Nuts3 level (provinces), exhibits similar patterns in neighboring provinces.

Formally, the index is defined as:

$$I = \frac{\sum_i^n \sum_j^n w_{ij} z_{x,i} z_{x,j}}{\sum_i^n z_{x,i}^2} \quad (1)$$

where:

- $z_{x,i} = X_i - M_x$  is the centered value of  $X_i$  in the  $i$ -th territorial unit
- $w_{ij}$  are the elements of the spatial contiguity matrix  $W$  row-normalized.

The interpretation of Moran's  $I$  is as follows:

- $> -\frac{1}{n-1} \rightarrow I$  positive spatial autocorrelation (clustering)
- $= -\frac{1}{n-1} \rightarrow I$  random distribution
- $< -\frac{1}{n-1} \rightarrow I$  negative spatial autocorrelation (dispersion)

Positive autocorrelation suggests that neighboring provinces tend to exhibit similar values, for instance, areas with smart specialization or high levels of digitalization clustering together, whereas negative autocorrelation points to a repulsive pattern.

### 3.2.2 Global bivariate Moran's I

The bivariate Moran's I measures the spatial association between two different statistical variables, X and Y. Specifically, it assesses whether the values of variable X observed in a province are associated with the values of variable Y observed in geographically contiguous provinces.

Formally:

$$I_b = \frac{\sum_i^n \sum_j^n w_{ij} z_{x,i} z_{x,j}}{\sum_i^n z_{x,i}^2} \quad (2)$$

where:

- $z_{x,i} = X_i - M_x$  and  $z_{y,j} = y_j - M_y$ ;
- $w_{ij}$  as above

In the classical formulation by Wartenberg (1985), the bivariate Moran's I takes a value of zero in the absence of spatial association between the variables, whereas positive or negative values indicate, respectively, clustering or dispersive processes. In other words, in the classical bivariate version (Wartenberg, 1985), the index equals zero when there is no autocorrelation, and it is positive or negative in the presence of, respectively, aggregative or dispersive spatial patterns.

The two local indices, obtained as an algebraic decomposition of the global index, identify the territorial units that contribute significantly to the formation of global autocorrelation.

The univariate LISA statistic can be expressed as follows:

$$L_i = z_{x,i} \sum_j^n w_{ij} z_{x,j} \quad (3)$$

It identifies clusters of territorial units as follows:

- $L_i > 0$  indicates HH or LL clusters, composed of units with high or low values of X;
- $L_i < 0$  indicates HL or LH clusters, composed of units with high values of X contiguous to units with low values of the same variable.

The bivariate LISA statistic accounts for the spatial configuration of the correlation between X and Y, highlighting the territorial units in which this correlation is statistically significant.

This statistic can be expressed as follows:

$$L_{bi} = z_{x,i} \sum_j^n w_{ij} z_{y,j} \quad (4)$$

- $L_{bi} > 0$ : it indicates a positive correlation, where high (or low) values of variable  $X$  in unit  $i$  are associated with high (or low) values of  $Y$  in the neighboring territorial units.
- $L_{bi} < 0$ : it indicates a negative correlation, where high values of variable  $X$  are associated with low values of variable  $Y$  in neighboring units, or vice versa.

According to the statistics  $L_i$ , the provinces can be clustered into four groups, in addition to the cluster of provinces not giving a significant contribution to global autocorrelation and to the cluster of provinces not confining with other provinces (minor islands):

1. High-High (HH) - contiguous units with values of the variable  $x$  higher than the mean;
2. Low-Low (LL) - contiguous units with values of the variable  $x$  lower than the mean;
3. High-Low (HL) - units with values of the variable  $x$  higher than their contiguous units;
4. Low-High (LH) - units with values of the variable  $x$  lower than their contiguous units.

The first two clusters are composed of homogeneous units (referring to the variable of interest), while the last two clusters concern abnormal cases or enclaves.

#### 4. Data and indicators

The information used to construct the S3 classification and related indicators was collected through the 2022 Istat business census, which included a specific section of the questionnaire dedicated to smart specialization. The census was carried out in 2022 (reference years 2020-2022), and covered around 280,000 enterprises with more than three employees. This sample represents about 24% of Italian enterprises and accounts for 84% of the national value added, 77% of total employment (12.7 million workers), and 91% of employees.

The survey provides a detailed picture of the structure of the Italian economic system and contributes to updating and improving the quality of business statistical registers. The census is based on a new statistical production process that integrates several existing data sources, including the Statistical Business Register (SBR) and the Frame SBS (Structural Business Statistics), an integrated system of administrative and statistical data produced annually by Istat to estimate firms' economic performance.

In addition, thematic registers are incorporated, such as the employment register, which contains detailed information on employment in active enterprises and represents the core of the new employment information system, structured as a Linked Employer–

Employee Database (LEED). Among the basic indicators used to compute the S3 index score is, for instance, the share of employees with tertiary education.

The indicators used in this work contribute to the construction of the innovation composite index, one of the five indices of the S3 framework<sup>4</sup>. In particular, to capture the two dimensions of the “twin transition” two specific indicators are used:

- 1) *Digit* (2.3 - Innovation in technology and digital fields). It represents the level of digitalization of enterprises and is based on five elementary indicators (1. Intensity of investment in digitization; 2. Software acquisition/development; 3. Acquisition of hardware/network equipment; 4. Production of KETs (Key Enabling Technologies); 5. Utilization of KETs). The composite index varies between 0 (lowest value) and 1 (highest value).
- 2) *Green* (2.4 - Environmental sustainability). It represents the propensity of enterprises to invest in environmental sustainability and is based on ten elementary indicators (1. Intensity of investment in environmental responsibility; 2. Measures to improve occupational well-being; 3. Collective interest initiatives; 4. Investment in the installation of power generation facilities; 5. Wastewater treatment; 6. Containment of water withdrawals and consumption; 7. Involvement of suppliers attentive to environmental impact; 8. Redesigning the production process for environmental sustainability; 9. Drafting environmental/sustainability reports; 10. Acquire voluntary process/product certifications. The composite index varies between 0 (lowest value) and 1 (highest value)).

The analyses presented in this paper also make use of the composite index developed to monitor potential S3 enterprises:

*S3 index score*: composite index based on 5 indicators (1. Research and development; 2. Innovation, 3. Human capital; 4. the ability to foster local development; 5. Economic performances. The composite index varies from 90 (lowest value) to 130 (highest value).

#### **4.1 Preliminary analyses**

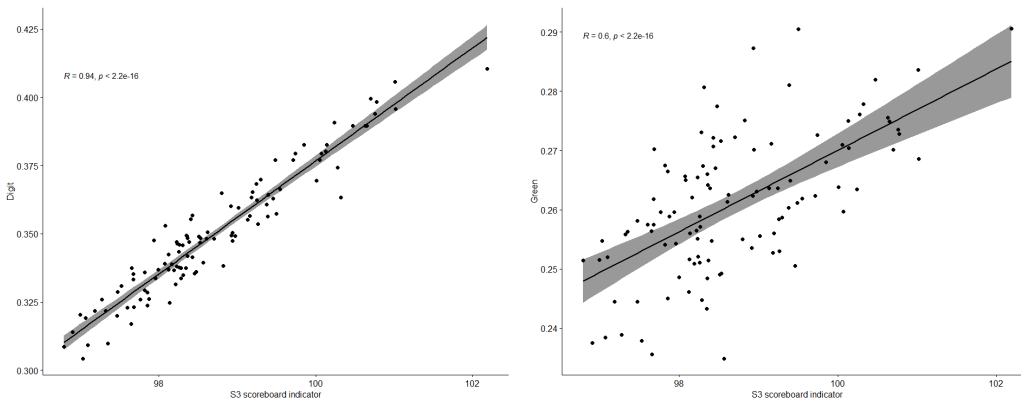
The data source and the indices defined for the S3 classification make it possible to assess the state of the twin transition at a detailed and granular territorial level, namely at the Nuts3 level. Preliminary analyses reveal groups of territories characterized by similarities and associations between levels of digitalization and sustainability.

---

<sup>4</sup> In the appendix, Table 1 illustrates the structure of the 5 composite indices, the related sub-dimensions, the elementary indicators (variables) and the methodologies used for aggregation.

The joint analysis of provincial average values of the digital and green indices together with the provincial S3 index is particularly informative. In particular, a stronger relationship between digitalization and smart specialization can provide insights into potential policy interventions aimed at balancing sustainable development and technological advancement (Figure 2).

Figure 2. Correlation between S3 score index and digit (first graph), S3 index and green (second graph), 2020-2022

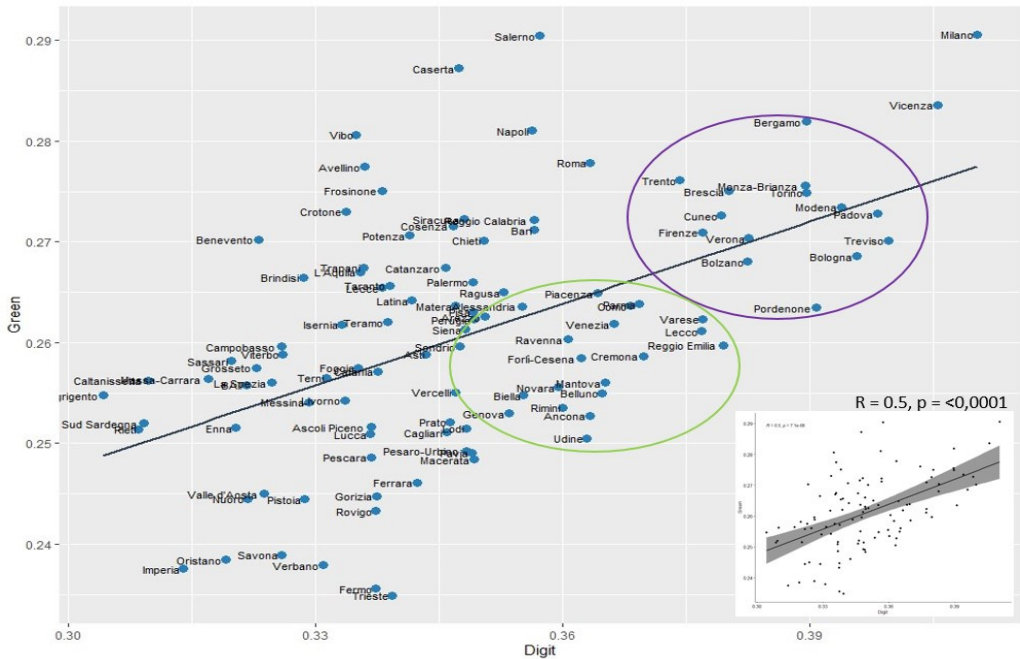


Source: our elaborations based on data from enterprises census and statistical business registers, 2022, Istat

An initial territorial analysis highlights a group of provinces characterized by high levels of economic development and a strong combination of digital intensity and environmental orientation. These territories are mainly located in north-eastern Lombardy and Veneto and also include provinces from other regions, such as Florence, Bologna and Cuneo (highlighted in the purple circle in Figure 3).

Another cluster (shown in green) displays significant associations between green and digital dimensions. This group includes provinces in north-western Piedmont, such as Biella and Novara, as well as several provinces along the northern Adriatic coast (Ravenna, Forlì-Cesena, Rimini and Ancona). More generally, some outlying territories in southern Italy exhibit high levels of environmental sustainability that are not matched by a comparable degree of digitalization.

Figure 3. Distribution of Italian provinces for digitalization (digit) and sustainability (green) index, Nuts3, 2022

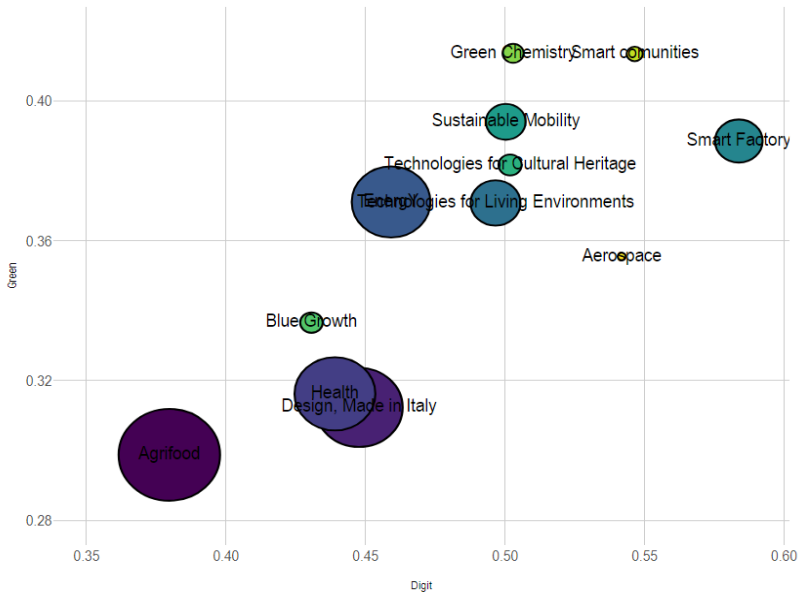


Source: our elaborations based on data from enterprises census and statistical business registers, 2022, Istat

The analysis presented in this study also includes an exploratory exercise that examines the state of the digital and green transitions across the twelve areas of specialization of the Italian smart specialization strategy described above. Figure 4 shows the relationship between the average values of the two composite indices, green and digital, across the twelve areas, with bubble size reflecting the number of enterprises in each area.

Clear differences emerge across areas of specialization in terms of the integration of green and digital dimensions. The smart factory area stands out, with above-average values on both indices. By contrast, the agrifood, blue growth, health, and design areas display similar patterns, combining relatively low levels of digitalization, particularly in the agrifood sector, with a stronger focus on environmental responsibility.

Figure 4. Distribution of Italian S3 areas for digitalization (digit) and sustainability (green) index, 2022



Source: our elaborations based on data from enterprises census and statistical business registers, 2022, Istat

## 5. Results

The entrepreneurial structure of the Italian territory shows marked differences in terms of production specializations concentrated in specific agglomerations, as well as in value added and average firm size. As a result, a purely descriptive territorial analysis may fail to capture similarities in production patterns or meaningful spatial connections.

For this reason, following the analysis of the territorial distribution at the Nuts 3 level, a spatial autocorrelation analysis was conducted using Moran's I and local indicators of spatial association (LISA) (Anselin, 1994).

### 5.1 Spatial correlation analysis of digitalization, sustainability and smart specialization indices

The univariate global Moran's I indicates significant spatial autocorrelation in the distribution of the S3 index score, the digitalization index and the of environmental sustainability index, with values of  $I = 0.3748$  ( $p < 0.01$ ),  $I = 0.2085$  ( $p < 0.01$ ) and  $I = 0.2075$  ( $p < 0.01$ ).

The S3 index shows the highest level of autocorrelation, pointing to an aggregative spatial pattern in which neighboring provinces tend to display similar values. For the twin transition indices, spatial autocorrelation is lower but still statistically significant, suggesting that their spatial distribution is not random.

To identify the provinces that contribute most to the global autocorrelation, the analysis was extended to the local level. This decomposition, carried out using Local Indicators of Spatial Association (LISA), makes it possible to distinguish provinces with no statistically significant contribution (shown in white on the maps) from those forming four distinct clusters that together shape the global Moran's I.

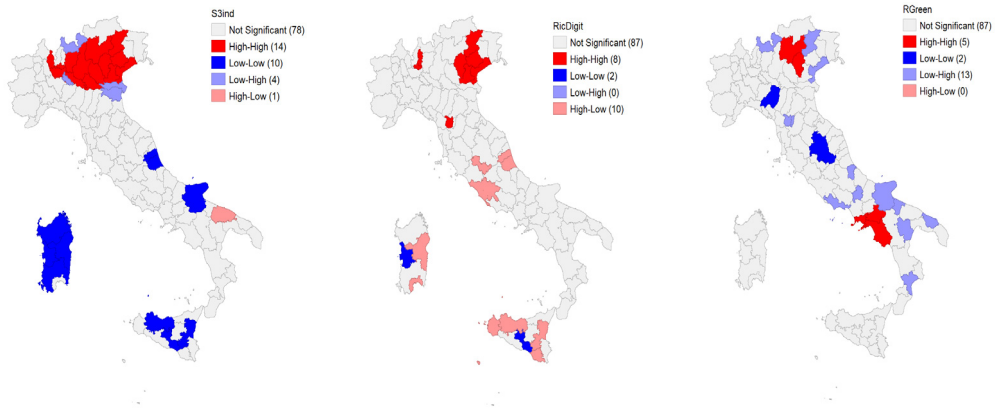
The first cluster identified is noteworthy, being the only one composed of contiguous provinces with an S3 score above the average of other provinces (High-High, HH, shown in red on the map). It includes provinces in the north-eastern regions of Lombardy (all provinces except Como, Lecco, Pavia, Lodi, and Sondrio), Veneto (except Rovigo), and the autonomous province of Trento. This cluster thus represents an area of strong propagation of smart specialization levels; however, some neighboring provinces, Sondrio, Lodi, Rovigo, and Ferrara, remain at lower levels (Low-High, LH, shown in violet), indicating limited spillover effects.

A second group (Low-Low, LL, in blue) of contiguous provinces is mainly located in Sardinia and Sicily, characterized by uniformly low S3 scores. The province of Bari represents an interesting exception: the only southern province with a high S3 score, forming a highly specialized enclave in a region otherwise dominated by low-specialization clusters, particularly in Sardinia, Sicily, and the area spanning Abruzzo to Puglia.

The spatial pattern of provinces according to the digitalization index appears simpler. Here, the HH cluster is limited to Pordenone, Padua, Treviso, Venice, and Vicenza. More widespread are HL clusters, composed of highly digitalized provinces adjacent to provinces with low index values, forming digitalization enclaves mainly in Sicily, Sardinia, and central Italy.

Greater territorial complexity emerges with the green index, which exhibits all four clusters identified by the LISA analysis. This indicates significant geographical heterogeneity beyond the traditional north-south divide. Specifically, the two HH clusters are located in north-eastern Veneto and southern Campania, while LL and LH clusters are distributed across the country from north to south, lacking a clear geographical pattern.

Figure 5. Local Moran index (LISA) with S3 index score (first map), digit (second map) and green index (third map)



Source: our elaborations based on data from enterprises census and statistical business registers, 2022, Istat

## 5.2 Multiple spatial correlation between twin transition and S3

The multivariate version of Moran's autocorrelation index measures the spatial correlation between different indicators observed in contiguous provinces. It therefore allows an assessment of whether, and to what extent, the spatial structure facilitates spillover effects across distinct dimensions of economic innovation. In this study, the systematic spatial association among the S3, digital, and green indices is investigated.

When interpreting the results of the bivariate Moran's I statistic (Table 1), it should be noted that the first label, preceded by "W", refers to the indicator measured in the provinces adjacent to the  $i$ -th province. As discussed above, the two indicators considered in each bivariate specification capture different dimensions of innovation.

The strongest spillover effects emerge for the WS3Ind/RDigit and WRDigit/S3Ind specifications, indicating a substantial and mutually reinforcing relationship between these two dimensions. In other words, spatial proximity exerts an almost symmetrical influence between the digitalization levels of the  $i$ -th province and the S3 index values of neighboring provinces. Similarly, higher levels of S3 specialization in the  $i$ -th province positively affect the degree of digitalization in adjacent provinces. By contrast, the bivariate correlations involving the remaining indicators are less pronounced, although they remain statistically significant and largely symmetrical.

*Table 1. Multiple spatial correlation index. 2022*

<i>S3, Digit, Green index</i>	<i>I (Moran)</i>	<i>pvalue</i>
WRDigit/S3Ind	0.259	<0.005
WRGreen/S3Ind	0.118	<0.05
WS3Ind/RDigit	0.261	<0.01
WRGreen/RDigit	0.103	<0.005
WRS3Ind/RGreen	0.101	<0.05
WRDigit/RGreen	0.109	<0.01

Source: our elaborations based on data from enterprises census and statistical business registers, 2022, Istat

The bivariate Local Moran statistic makes it possible to delineate an initial geography of innovation based on the interaction between the spatial correlation of different indicator variables and the provincial spatial structure. Of particular interest is the HH cluster for the WS3Ind/RDigit specification, located between Lombardy and the north-east. Similar territorial dynamics are observed for the WRDigit/S3Ind indicators; in this case, the presence of HL areas contiguous to HH clusters is noteworthy, as such spatial proximity may signal a potential expansion of high–high clusters.

Moreover, this second spatial configuration highlights the traditional geo-economic divide of the Italian peninsula, with highly specialized areas concentrated in the north and lower levels of specialization prevailing in the south. Finally, the HH clusters associated with the twin transition, identified both in the north-east and in parts of central-southern Italy, are particularly relevant, as is the presence of HL areas in northern Sicily, in several provinces of central Italy, and in Liguria, suggesting heterogeneous and spatially fragmented pathways toward digital and green transformation.

### ***5.3 Spatial analysis of smart specialization areas***

The Italian smart specialization areas described above are analyzed here from a territorial perspective in order to identify spatial correlations associated with the twin transition. The univariate Moran's I was computed for each indicator across the twelve areas of specialization. The results of this preliminary analysis show that Moran's I is statistically significant in only nine areas of specialization, and primarily for the indicator capturing the number of firms located in each province. By contrast, the spatial autocorrelation values for the digitalization and sustainability indicators across the different thematic areas are not statistically significant (Table2).

The design and made in Italy area of specialization exhibits the highest level of spatial autocorrelation, followed by smart factory and green chemistry, with respect to the distribution of production units across Italian provinces. However, despite being statistically significant, these correlation values remain relatively low, suggesting that the identified areas of specialization do not display a strong territorially distributed linkage across provinces. With regard to the digitalization index, significant spatial autocorrelation is observed only for the aerospace and blue growth areas of specialization.

Table 2. Spatial correlation index for the twelve S3 areas. 2022

<i>Specialization areas</i>	<i>Enterprise</i>	<i>S3 index</i>	<i>Digit index</i>	<i>Green index</i>	<i>Moran (I)</i>
Design, Made in Italy	15717	102.6	0.447	0.312	0.3017
Smart Factory	4929	107.6	0.583	0.388	0.2480
Green Chemistry	1762	106.0	0.502	0.413	0.2256
Agrifood	21250	100.5	0.379	0.298	0.1653
Technologies for Living Environments	5358	104.3	0.496	0.37	0.1484
Energy	12721	104.3	0.459	0.371	0.1228
Aerospace	1379	107.2	0.541	0.355	0.1749
Blue Growth	1857	102.4	0.43	0.336	0.1130
Sustainable Mobility	3656	105.5	0.5	0.394	-
Health	13447	102.0	0.439	0.316	-
Smart communities	1524	106.8	0.546	0.413	-
Technologies for Cultural Heritage	1897	105.2	0.501	0.381	-

Source: our elaborations based on data from enterprises census and statistical business registers, 2022, Istat

## 6. Conclusion

The innovative interpretation of smart specialization, as embedded in the evolution of place-based development policies, moves beyond a traditional focus on the concentration of resources in the most competitive sectors. Instead, in line with the RIS3 framework and insights from evolutionary economic geography, it emphasizes the strategic concentration of resources in activities with growth potential that are rooted in territorially specific capabilities and knowledge bases. This policy trajectory can stimulate firms' innovative dynamism by facilitating diversification into new products and services that are related to existing regional specializations. The intersectoral relationships that emerge through this process highlight the central role of related variety at the local level as a key mechanism underpinning regional innovation and path development.

The analysis of relationships among areas of specialization has made it possible to identify and assess the potential for territorial cooperation, as well as the presence of research networks that may influence the technological performance and economic growth of neighboring provinces.

The results of this study enable an assessment of the investment priority choices outlined in regional smart specialization strategies by comparing them with the evidence emerging from spatial correlation analyses. Areas of specialization are defined by moving beyond the conventional boundaries of economic activities, identifying a critical mass of innovative firms through the five S3 dimensions, as measured by purpose-built composite indicators.

This contribution highlights the relevance of the spatial dimension for innovation, as the introduction of new products and production processes entails interactions among a plurality of actors, such as universities and research centres, which are facilitated by territorial proximity. The circulation of information and knowledge (spillover effects) generates positive externalities that also benefit actors who have not directly contributed to knowledge creation but operate in close spatial proximity to its source.

From a policy perspective, the territorial dimension emerges as an active and strategic variable rather than a merely descriptive one, made operational through the geocoding of data at the Nuts3 level. The findings of this study provide several insights that are relevant for the evaluation and refinement of regional smart specialization strategies (RIS3 - Research and Innovation Strategies for Smart Specialization).

First, the results confirm the suitability of the proposed composite indicators in capturing key dimensions of innovation, namely digitalization and environmental sustainability, from a territorial perspective, supporting their use as monitoring and evaluation tools for S3 implementation. Second, the evidence of spatial spillovers underscores the importance of territorial contiguity among areas, suggesting that RIS3 policy design should explicitly account for interregional interactions and functional economic areas.

Third, the application of spatial analysis techniques to composite indicators identifying S3 firms enables the delineation of a more nuanced geography of innovation, incorporating qualitative and strategic dimensions of firm behavior. These results move beyond the conventional North–South divide and provide a more articulated interpretation of Italian regional development in terms of r&d and innovation propensity, as well as the relationship between digitalization and sustainability.

Fourth, while the spatial concentration of specialized and highly digitalized territories largely reflects the historical North–South divide, the presence of sustainability-oriented dynamics in several southern territories, often not matched by equivalent levels of digitalization, highlights the need for differentiated and context-sensitive policy interventions.

Fifth, the experimental territorial analysis of the twelve areas of specialization with respect to the twin transition reveals the absence of clearly defined territorial clusters that simultaneously integrate digital and sustainability dimensions. This finding points to a potential fragmentation of policy priorities and suggests the need for stronger policy coordination and integration across S3 domains.

Finally, the limited spatial correlation observed among the twelve areas of specialization may indicate a low degree of territorially embedded technological and innovative related variety, which could constrain processes of regional diversification and path development. From a policy standpoint, this calls for targeted measures aimed at strengthening cross-sectoral linkages, fostering knowledge recombination, and enhancing complementarities between existing regional capabilities within the S3 framework.

With regard to the proposed results, it is important to account for potential distortions arising from the territorial level of data granularity. Within the information set used to construct the composite indicators, the main data source is the business census, which is based on a sampling design at the Nuts2 level and, for selected domains, at the Nuts3 level. Consequently, the delineation of the twelve areas of specialization, particularly in regions characterized by a low density of smart specialization firms, may rely on a limited critical mass of enterprises that is potentially not fully representative at the area level.

The new classification of S3 firms incorporates the dimensions underlying the smart specialization strategy, comprising five main dimensions and 35 sub-dimensions, into the composite index. This structure may pose potential endogeneity issues when evaluating the contribution of certain indicators to territorial development. However, the possibility of scoring each individual dimension allows for an assessment of the contribution of these factors and makes it possible to identify which territories exhibit strengths or shortcomings in supporting territorial development.

Regarding policy implications, an immediate application of the proposed results concerns regional administrations, as well as national authorities, in the planning and programming of S3 funding. Funding allocations can be differentiated not only by territory but also by thematic objectives, based on identified gaps, for example, in digitalization or sustainability.

The contribution of this work to supporting the monitoring and evaluation of innovation policies extends beyond guidance on balancing productivity, innovation, and environmental sustainability. It also encompasses the territorial dimension of well-being. This aspect, emphasized by the S3 strategy, is captured through indicators measuring improvements in job quality and the activation of collective-interest initiatives by firms.

This work provides a novel contribution to smart specialization policy by applying spatial analysis to a new and highly informative dataset that integrates qualitative cen-

sus data with quantitative administrative information. This approach fills an important informational gap and delivers insights at a detailed territorial level. The findings reinforce the S3 policy focus on place-based territorial strengths and the development of distributed and fragmented knowledge within regions. They confirm existing development trajectories identified in the literature and, importantly, highlight territories with potential growth pathways, offering actionable guidance for targeted investment, funding allocation, and regional innovation planning.

## References

- ASCANI, A., BALLAND, P.-A., & MORRISON, A. (2020). Heterogeneous networks and regional resilience, *Journal of Economic Geography*.
- ANSELIN, L. (1995). Local Indicators of Spatial Association-LISA. *Geographical Analysis*, 27(2), 93-115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338>
- BALLAND, P.A., BOSCHMA, R., CRESPO, J., & RIGBY, D.L. (2019). Smart specialization policy in the EU: relatedness, knowledge complexity and regional diversification, *Regional Studies*, 53:9, 1252-1268.
- BARCA, F. (2009). Agenda for a Reformed Cohesion Policy; a place-based approach to meeting European Union challenges and expectations, *European Communities*.
- BINZ, C., & CASTALDI, C. (2024). Toward a normative turn in research on the geography of innovation? Evolving perspectives on innovation, institutions, and human well-being, *Progress in Economic Geography*, Volume 2, Issue 2.
- BOSCHMA, R. (2005). Rethinking regional innovation policy. In: Fuchs, G., Shapira, P. (eds) *Rethinking Regional Innovation and Change*. Economics of Science, Technology and Innovation, vol 30. Springer, New York, NY. [https://doi.org/10.1007/0-387-23002-5\\_12](https://doi.org/10.1007/0-387-23002-5_12)
- BOSCHMA, R. (2005). Proximity and Innovation: A Critical Assessment. *Regional Studies*, 39(1), 61-74. <https://doi.org/10.1080/0034340052000320887>
- BOSCHMA, R. (2017). Relatedness as driver of regional diversification: a research agenda. *Regional Studies*, 51(3), 351-364. <https://doi.org/10.1080/00343404.2016.1254767> Losacker, S., Hansmeier, H., Horbach, J. et al. (2023). The geography of environmental innovation: a critical review and agenda for future research. *Rev Reg Res*. <https://doi.org/10.1007/s10037-023-00193-6>.
- D'ADDA, D., IACOBUCCI D., & PALLONI R., (2020). Relatedness in the implementation of Smart Specialisation Strategy: a first empirical assessment. *Regional Science*. 99:405-425.
- FORAY, D., & GOENAGA, G. (2013). "The goals of Smart Specialisation," JRC Research Reports JRC82213, Joint Research Centre
- FORAY, D. (2015). Smart Specialisation. Opportunities and Challenges for Regional Innovation

Policy. London: *Routledge*.

- ISTAT (2022). Indicatori per la strategia di specializzazione intelligente. <https://www.istat.it/it/archivio/273548>
- MAZZIOTTA, M., & PARETO, A. (2016). On a generalized non-compensatory composite index for measuring socio-economic phenomena. *Social Indicators Research*, 127, 983-1003.
- MIEDZINSKI, M., CIAMPI STANCOVA, K., MATUSIAK, M., & COENEN, L. (2021). Addressing sustainability challenges and Sustainable Development Goals via Smart Specialisation. Towards a theoretical and conceptual framework, *Publications Office of the European Union*, Luxembourg, ISBN 978-92-76-42380-5, doi:10.2760/410983, JRC126448.
- RAKHMATULLIN, R., & HEGYI F.B. (2021). Exploring the potential of thematic Smart Specialisation Partnerships to contribute to SDGs. *Publications Office of the European Union*, Luxembourg, ISBN 978-92-76-41059-1, doi:10.2760/728381, JRC125210.
- SALAMONE, S., & ARMENISE M., (2022). Come valutare la coerenza delle aree di specializzazione intelligente regionali attraverso un indice di complessità economica. *Regional Economy*, vol. 6(Q3), pages 22-31.

## APPENDIX

Figure A1

DIMENSIONS	SUB DIMENSIONS	INDICATORS	VARIABLES
1. Research & Development	1.1 Ability to expand knowledge assets and innovation in production	R&D expenditure	1. Intensity of R&D investment 2. Acquisition of R&D services
	1.2 Applications and experimentation in R&D	R&D facilities	1. In-house R&D activities
		Output	1. Acquisition of licenses and patents
2. Innovation	2.1 Product and process innovation	Product and process innovation	1. Technical and aesthetic design 2. Acquisition of machinery/equipment for innovations
	2.2 Organizational and marketing innovation, innovative finance	Organizational and marketing innovation, innovative finance	1. Marketing for launching new goods/services 2. Sources of innovative financing of the enterprise
	2.3 Innovation in technology and digital fields	Digitalization	1. Intensity of investment in digitization 2. Software acquisition/development 3. Acquisition of hardware/network equipment 4. Production of KETs 5. Utilization of KETs
	2.4 Environmental and social sustainability	Environmental and social sustainability	1. Intensity of investment in environmental and social responsibility 2. Measures to improve occupational well-being 3. Collective interest initiatives 4. Investment in installation of power generation facilities 5. Wastewater treatment 6. Containment of water withdrawals and consumption 7. Involvement of suppliers attentive to environmental impact 8. Redesigning the production process for environmental sustainability 9. Drafting environmental/sustainability reports 10. Acquire voluntary process/product certifications
3. Human capital	3.1 Quality of human resources and skills development	Training and skills (level)	1. Training activities other than mandatory training 2. Intensity of investment in human capital and training 3. Employees with tertiary education 4. Staff training on innovations
4. Capacity to drive the development of territories	4.1 Ability to create and expand relationships to increase innovative capacity	Relationship networks	1. Formal agreements with universities/research centers 2. Formal agreements with public administration 3. Formal agreements with enterprises 4. Formal agreements for the development of new products/processes 5. Formal agreements for access to new skills/technologies
	4.2 Openness to new markets and internationalization	Markets and internationalization	1. Intensity of investment in internationalization 2. Formal agreements for access to new markets/customers
5. Economic performance	5.1 Productivity levels	Productivity levels	1. Value added per employee

Source: our elaborations based on data from enterprises census and statistical business registers, 2022, Istat



## **THE ROLE OF ARTIFICIAL INTELLIGENCE IN SUPPORTING HUMAN DECISIONS: DIGITAL TWIN FOR STREAMLINING MARKET RESEARCH**

### **IL RUOLO DELL'INTELLIGENZA ARTIFICIALE NEL SUPPORTARE LE DECISIONI UMANE: LE DIGITAL TWIN PER SEMPLIFICARE LE RICERCHE DI MERCATO**

*Lino Coscione<sup>1</sup>, Cinzia Paterlini<sup>2</sup>, Salvatore Fagnoli<sup>3</sup>*

#### **Sommario**

Il presente articolo analizza il ruolo dell'Intelligenza Artificiale (IA) nell'ottimizzazione dei processi decisionali umani, con un focus sull'impiego della tecnologia dei Digital Twin nella ricerca di mercato. Utilizzando il framework TP Infinity TWIN come caso studio, la ricerca illustra come l'IA possa generare gemelli digitali di consumatori per testare scenari qualitativi e quantitativi. Attraverso l'integrazione di input provenienti da molteplici fonti – tra cui dati di sondaggi, *social listening* e tecniche di *Natural Language Processing* (NLP) – vengono costruiti profili realistici testati mediante focus group virtuali. Le osservazioni empiriche suggeriscono progressi significativi nella scalabilità dei dati, nella velocità e nella generazione di *insight*. I risultati indicano, tuttavia, che tali modelli computazionali operano come un livello secondario di supporto alle decisioni, fungendo da complemento tecnico – piuttosto che da sostituto totale – dei metodi tradizionali. All'interno di questo framework, infatti, la supervisione umana funge da fondamentale salvaguardia epistemologica, garantendo la validità interpretativa e mitigando i rischi inerenti ai bias algoritmici.

#### **Abstract**

*This article investigates the role of Artificial Intelligence (AI) in enhancing human decision-making, with a focus on the use of Digital Twin technology in market research. Utilizing the TP Infinity TWIN framework as a case study, the study illustrates how AI-generated customer twins can simulate and test qualitative and quantitative scenarios. By integrating multi-source inputs, including survey data, social listening, and NLP techniques, realistic profiles are built and tested through virtual focus groups. Empirical observations suggest significant advancements in data scalability, speed, and insights generation. The findings indicate, however, that these computa-*

<sup>1</sup> TP Infinity, Rome, Italy - e-mail: lino.coscione@teleperformance.com

<sup>2</sup> TP Infinity, Milan, Italy - e-mail: cinzia.paterlini@teleperformance.com

<sup>3</sup> TP Infinity, Rome, Italy - e-mail: salvatore.fagnoli@teleperformance.com

*tional models function as a secondary decision support layer, acting as a technical complement to – rather than a total replacement for – traditional methods. Within this framework, human oversight serves as a critical epistemological safeguard, ensuring interpretative validity and mitigating the inherent risk of algorithmic bias.*

**Parole chiave:** Intelligenza artificiale, digital twin, ricerca di mercato, processi decisionali, dati qualitativi densi

**Keywords:** *Artificial intelligence, digital twins, market research, decision-making, thick data.*

Classificazione JEL: M31, C81, D83

## 1. Introduction

Artificial intelligence (AI) is increasingly being used to support human decision-making, especially in fields such as market research, where vast data and complex consumer behaviors must be understood. One increasingly relevant approach is the use of digital twin technology – creating virtual models of consumers or markets – to simulate scenarios and gather insights before taking real-world actions. By leveraging AI-driven simulations of these Digital Twins, companies can evaluate product ideas, marketing campaigns, and strategies in a controlled virtual environment. This article offers a conceptual and empirical analysis of the evolving Digital Twin landscape in market research. It first establishes a theoretical framework for AI-supported decision-making, clarifying the intersection between digital proxy architectures and market research. Subsequently, the TP Infinity TWIN model serves as a case study to illustrate recent developments in the application of synthetic modeling in consumer behavior analysis, iterative product testing, and predictive campaign simulation. Selected applications and empirical industry cases are examined in conjunction with relevant academic publications to provide a grounded perspective on the field. The concluding section provides a critical assessment of the operational benefits, limitations, and future evolutions of digital twins and AI-driven simulation in market research.

### 1.1 Definitions and Key Concepts

A **Digital Twin** is a virtual replica of a real-world entity – this could be a physical object, system, person, or process – that is synchronized with real data to mirror the original's state and behavior. Traditional uses of digital twins in manufacturing and engineering allow companies to simulate real situations and outcomes – such as testing machine performance under various conditions – which ultimately helps them make better decisions with fewer risks. Key characteristics of Digital Twins include real-time

data updates (reflecting changes in the original system) and the ability to run simulations or “what-if” analyses.

In recent years, this concept has expanded beyond machines to model human-centered systems. Companies and researchers are now exploring digital twins of customers and markets – essentially virtual consumer avatars or simulated marketplaces – as a tool for business insight. Gartner (2022, p. 2) describes a Digital Twin of a Customer (DToC) as building on the traditional marketing persona concept to “provide context and predictions of future consumer behavior”, relying on data from both online and physical interactions to simulate a customer’s experience in real time.

Essentially, a Digital Twin within the marketing domain is a data-driven virtual model of a consumer (or consumer segment) that simulates behavior, latent preferences, and decision patterns.

**AI-supported decision-making** refers to the use of artificial intelligence algorithms and systems to augment human decision processes. Rather than replacing human judgment, AI acts as a decision aide analyzing large data sets, identifying patterns, forecasting outcomes, and even recommending optimal actions. The goal is to help humans make faster, more informed, and evidence-based decisions.

In the market research realm, this means AI systems can process consumer research data (surveys, social media, sales figures, etc.) to find trends and then simulate decisions (“What happens if we raise the price? If we target Segment X with this campaign?”) before a human manager chooses a course of action. Importantly, AI-supported decisions can be categorized as either fully automated systems (the AI decides within set parameters) or human-in-the-loop configurations (AI provides options or analysis that a human decision-maker considers). Within the domains of marketing and consumer insights, the latter is more common: while computational architectures manage large-scale data processing and model execution, humans set objectives and interpret the results, ensuring that strategic choices align with business context and ethics.

**AI-powered Digital Twins** offer a way to streamline these processes through simulation. In this approach, companies create Digital Twins of their consumers or target market segments (sometimes called “synthetic consumers” or “virtual customers”) and use AI algorithms to simulate how these virtual individuals would behave under different conditions. Essentially, this merges rich data (the foundation of a digital twin) with advanced modeling (provided by AI) to create a virtual tester for decision-making.

This approach allows organizations to gain consumer insights and test strategies virtually before committing resources in real-world settings, thereby streamlining the research process and reducing uncertainty.

Recent advances in AI and data availability have accelerated the use of Digital Twins in market research. Notably, the rise of **generative AI and large language models**

(LLMs) has made it feasible to create highly detailed and dynamic consumer simulations.

### ***1.2 AI in market research***

Making a decision, especially in a strategic or professional context, requires consideration of several key elements that help ensure that the choice is well thought out, informed, and aligned with the intended objectives. Several factors contribute to decision-making processes: contextual knowledge, supporting data collection, analysis, risk and consequence assessment, and available resources. The collection and analysis of elements are inevitably determined by the human factor, but AI can play a significant supporting role in improving the quality, speed, and effectiveness of decisionmaking.

The introduction of artificial intelligence into business and research processes should not be seen merely as a technological upgrade, but rather as a strategic change that requires careful planning. It is critical that it be a conscious integration: adopting AI in a way that supports and enhances human capabilities, without replacing them, to avoid losing human critical vision, intuition, and creativity, elements that remain irreplaceable for interpreting data, making informed decisions, and dealing with the complexities of the real world. On the contrary, conscious integration of AI brings high efficiency in decision-making processes, significant cost reduction, and faster access to information, analytical outputs, and hypotheses.

While AI is powerful in collecting and analyzing massive amounts of data, it is human intelligence that drives the design, evaluation, and continuous improvement of these models. Human input is crucial not only at the final decision-making stage, but throughout the entire decision-making process: bringing contextual knowledge, ethical values, and practical experience that AI cannot replicate. In addition, human feedback is essential to refine the algorithms, ensuring that they remain relevant and accurate over time.

To date, **Artificial Intelligence is already widely applied in market research in various fields:**

- Collection stage: from gathering data from the Web (Web scraping) or Social Media (social listening), to recognizing and analyzing visual content (images and videos) so as to recognize and understand consumer preferences;
- Analysis stage: text mining (extracting information from text), identifying trends and sentiment analysis, defining preference patterns from visual content, to predicting future consumer behavior (predictive analytics);
- In addition, AI is widely used in the creation of reports and dashboards with interactive visualization.

These applications may generate several **operational benefits for decision-makers**:

- Speed: analysis of large volumes of data quickly;
- Accuracy: reducing human error and improving the precision of analysis;
- Scalability: adaptability to projects of any size;
- Deeper insights: discovering hidden patterns and correlations in the data;
- Personalization: creating customer-centric experiences.

But the the **following limitations** should not be ignored:

- Data quality: the quality of analysis depends on the quality of data input;
- Privacy: personal data management and compliance with privacy regulations;
- Model interpretability: difficulties in interpreting more complex deep learning models;
- Cost: high initial investment in implementing AI solutions.

## 2. Methodological framework and case analysis: the TP Infinity approach

The proposed framework serves as a functional exemplar of human-AI collaboration in practice. This model illustrates how the integration of synthetic intelligence can optimize operational efficiency and predictive precision while maintaining the rigor and ethical integrity of the decision-making process.

### 2.1 Digital Twins

**TP Infinity**, the specialized consulting and analytics division of TP, has developed a methodology centred on **Digital Twins**: a framework for constructing synthetic, data-enriched profiles that represent specific consumer segments. This architecture serves to optimize the market research process by facilitating the rapid generation of insights and the iterative development of concepts. By leveraging this synthetic approach, decision makers can access high-fidelity information with significantly reduced temporal and financial overhead, at low cost and with a good level of quality.

The methodology entails creating a synthetic profile, or Digital Twin, for each target consumer segment (also defined as Digital Persona or Icon), a synthetic sample that allows complex scenarios to be simulated and in-depth insights to be obtained without the logistical and time limitations of traditional research.

A central methodological contribution of Digital Twins is the deployment of advanced algorithms to construct **multidimensional customer profiles** through the integration of data from different sources. These profiles incorporate demographic and behavioral data from conventional qualitative and quantitative surveys as well as inputs such as psychographic dimensions, personal preferences and affective reactions derived from social listening. This analytical depth is facilitated by machine learning algorithms, ca-

pable of identifying correlations and latent patterns that frequently remain undetectable through standard non-automated analytical frameworks.

The methodology operates across multiple data layers: demographic profiles, behavioral segmentations, psychographic mapping, and real-time sentiment acquisition via social listening methodologies. Machine learning architectures, including transformers and recurrent neural networks, are trained on these multidimensional inputs to generate realistic responses to research stimuli. Each Digital Twin is constructed using clustering and pattern recognition and further refined using expert-driven validation to ensure realism and utility. Prompt engineering is a key aspect, ensuring that AI-generated answers remain contextually accurate and aligned with segment-specific logic. This methodological rigor ensures that digital twins serve as credible testbeds for market hypotheses.

The **deployment of Digital Twins within qualitative research** contexts has yielded significant heuristic insights, though these results remain epistemologically distinct from traditional focus group data. This divergence arises from the distinct ontological paradigms of the two methodologies: whereas artificial intelligence (AI) operates within a framework of logical consistency and rational coherence, human subjects are driven by emotions, impulses and contradictions – aspects that are an integral part of human interactions.

A crucial aspect that distinguishes the two methodologies is the current incapacity of AI to capture and analyze nonverbal language, such as tone of voice, silences, posture, and conversational pacing. These elements are crucial in conventional focus groups, where nonverbal behavior can reveal much about participants' emotions and motivations. The responses provided by AI-created Digital Twins are always consistent and free of contradictions, whereas humans, in contrast, tend to be emotional, contradict themselves and react impulsively.

Moreover, in human focus groups, the discursive dynamics are characterized by spontaneity and non-linear interactions: participants comment, interrupt, and criticize each other, contributing to a lively and multidimensional discussion. This spontaneity and inter-subjective fluidity are difficult to replicate in a synthetic environment, where Digital Twins may lack the same level of interactive complexity and social reactivity.

To mitigate these constraints, it is imperative that AI models undergo rigorous training, ensuring that Digital Twins are built and enriched with specific and detailed data. This data must necessarily be rooted in human sources, such as qualitative and quantitative surveys, as well as data collected from social media and the web.

If AI is not fed with rich and accurate information, the responses obtained tend to be generic, stereotypical, and based on common sense, thus limiting the value of the research. In extreme cases, AI “hallucinations” can occur, where the system generates answers that do not adhere to reality. To mitigate this risk, the systematic integration of data

related to socio-demographic variables, consumer habits, behavioral patterns, product and brand preferences, as well as lifestyles is fundamental to the construction of high-fidelity profiles. In this context, Digital Twins are a starting point for preliminary analysis rather than a definitive conclusion, as is typical in traditional qualitative research.

## 2.2 Applications

Once rigorously created and refined, Digital Twins serve as scalable analytical frameworks capable of being deployed across diverse research contexts and varying stages of the investigative process. For instance:

- **Interactive Desks:** Digital Twins can be queried interactively, as if consulting advanced and rich databases. This allows for quick and articulate answers to specific questions (consistent with the data on which the model was trained and profiles constructed), thus facilitating research and information gathering in reduced time;
- **Idea Generation:** Digital Twins can provide creative insights for developing new products or services, creating communication campaigns, and designing social strategies. This utility is significantly augmented when profiles are enriched with data from social listening, providing a dynamic representation of emergent audience trends and latent preferences.

Beyond qualitative simulation, the Digital Twins platform has demonstrated empirical efficacy in concept testing, A/B experimentation, and discursive optimization.

In several pilot projects, digital twins were deployed for the ex-ante evaluation of brand initiatives, facilitating the identification of communicative assets with the highest perceptual alignment. These synthetic Personas were dynamically updated with campaign-related stimuli and provided feedback that mirrored real-life consumer reactions. This capability aligns with recent advancements in ‘decision intelligence’ systems, where simulations inform strategic choices by presenting validated predictions to human decision-makers. In practice, Twin approach has been integrated with customer journey analytics to simulate reactions across different touchpoints, providing an end-to-end view of engagement potential.

## 3. Conclusion and discussion

Empirical evidence derived from TP Infinity’s Digital Twins framework, alongside current industry benchmarks, indicates the high potential for the integration of Digital Twins architectures within strategic decision-making. The synthesis of artificial intelligence and Digital Twin technology into market research and decision-making processes offers several distinct advantages:

- **Operational Velocity and Agility:** Digital Twin simulations dramatically accelerate the feedback loop within the market research cycle: marketers can iterate campaign concepts multiple times in a week, or product teams can refine a concept repeatedly before any physical prototype is built;
- **Resource Optimization and Cost Efficiency:** conventional research methodologies (focus groups, test market launches, extensive field surveys) entail substantial resource allocation. By integrating high-fidelity simulations, organizations can achieve significant cost mitigation. While initial investments required for robust Digital Twin development are noteworthy, the marginal cost of subsequent simulations is minimal. Additionally, virtualized testing can reduce the opportunity cost of time, accelerating the transition from data acquisition to actionable insight.
- **Personalization and Hyper-Targeting:** Digital Twin frameworks facilitate an unprecedented degree of strategic individualization. By moving beyond aggregate segmentation toward the maintenance of micro-segment or even individual-level models, organizations can simulate heterogeneous responses to diverse strategic interventions with high precision.

Although the integration of Digital Twins into the marketing paradigm offers significant potential, the framework is subject to substantial methodological constraints and challenges. A rigorous acknowledgement of these limitations is essential to maintain objectivity and to delineate clear trajectories for future research and optimization;

- **Data Quality and Integration:** the fidelity of a Digital Twin of a customer or market is contingent upon the quality of the underlying data. Creating a realistic consumer twin requires the synthesis of different data streams – including transactional histories, web analytics, CRM records, social media, market research data. The aggregation and harmonization of these multifaceted datasets present significant technical and methodological complexities;
- **Model Bias and Validity:** AI models, especially machine learning or generative AI architectures, are susceptible to the propagation of systemic biases into simulations. As noted earlier, when using LLMs to generate personas or responses, researchers observed a tendency to skew simulations, producing overly optimistic or stereotypical characters if uncalibrated. Therefore, building Digital Twins from market research data or consumer-related data is pivotal together with researchers' and marketers' interpretations. In other words, uncritical reliance on AI-simulated respondents may lead to misleading conclusions. Ensuring validity is a challenge: one must constantly compare and update the Digital Twin against real-world observation;

- **Complexity and Technical Proficiency:** the development and governance of Digital Twins for market research requires a synthesis of sophisticated competencies across data science, AI modeling, and domain expertise in consumer behavior. It is a technically complex undertaking;
- **Epistemological Trust and Operational Integration:** even with high-fidelity simulations, the translation of synthetic findings into actionable strategic actions remains complex. Managers may question the reliability of a virtual result. Organizations need to develop new decision-making protocols to integrate these AI insights, which include training teams to understand the strengths and limits of simulations;
- **Privacy and Ethical Concerns:** constructing a Digital Twin of customer raises questions about privacy and data use. Regulations like GDPR require transparency about automated profiling and decision-making. Deployment necessitates strict adherence to data privacy jurisdictions and evolving ethical frameworks. This is critical when synthetic outputs inform interventions affecting the treatment of customers;
- **Not a Complete Substitute for Reality:** finally, it must be said that no matter how detailed, a Digital Twin is ultimately a model – a simplification of the real world. Human behavior, in particular, has elements of randomness and emotion that are extremely hard to capture fully in code. There will always be unpredicted events (a viral social media moment, a sudden change in consumer sentiment due to a news event, etc.) that undermine predictive accuracy. Over-reliance on simulations could lead to a false sense of security. Market research using Digital Twins should complement, not entirely replace, direct methods of getting feedback from real customers.

In conclusion, although the application of Digital Twins in qualitative research cannot completely replace traditional research methods, such as focus groups, it offers a complementary approach that, when integrated with human data and robust methodologies, can be a powerful tool for improving the efficiency and effectiveness of market research. Digital Twins, properly constructed and updated, thus become a valuable resource not only for collecting data, but also for generating innovative ideas and supporting strategic decisions in an agile and flexible manner.

AI-powered Digital Twins are facilitating a paradigm shift in market research, enabling accelerated, resource-efficient, and evidence-based strategic decisions. The TP Infinity framework illustrates how human-centered AI, when governed by ethical and strategic oversight, can enhance the decision-making process without replacing the human elements of creativity, critical thinking, and nuanced emotional judgment.

Used properly, Digital Twins become not just tools for insight generation but engines of innovation and strategy, blending data-driven logic with human relevance. The journey is ongoing, and the most successful applications will be those where AI and human intelligence work in harmony.

As organizations prioritize operational agility and granular personalization, Digital Twins represent a blueprint for integrating AI-driven simulations with human expertise. These architectures enable the virtualized prototyping of consumer responses, facilitating resource conservation and the extraction of latent insights prior to market implementation. By combining TP Infinity methodology with contemporary advancements in artificial intelligence, this article illustrates how human-AI symbiosis can facilitate a transition from reactive market research toward a proactive, strategic modality. This collaborative intelligence model — characterized by machine-led simulation and human-centric validation — proposes a significant evolution in evidence-based decision-making frameworks.

#### 4. Future trends and outlook

Looking ahead, the role of AI and Digital Twins technology in decision-making and market research is poised to grow even further. Several **key future trends** are emerging:

- **Hyper-Personalized Simulations:** we are moving towards an era of extreme personalization in marketing, and Digital Twins will be central to that. By 2025 and beyond, marketers are expected to routinely run personalized campaign simulations targeted to very specific audiences – essentially creating a Digital Twin for each significant customer segment (or even individual) and simulating tailored content for them;
- **AI-Augmented Decision Intelligence Platforms:** we anticipate more integrated decision intelligence systems where Digital Twins are part of a larger AI platform advising business leaders. Instead of ad-hoc simulations run by data scientists, future systems (possibly AI assistants) could continuously analyze market conditions, run simulations, and feed recommendations to decision-makers in real time;
- **Continuous Real-Time Twins:** Digital Twins are expected to become more dynamic and real-time, blurring the line between simulations and live analytics; we will likely see real-time customer twins that update with every transaction or interaction;
- **Advanced Behavioral Models and Psychology Integration:** on the research front, digital twin models of humans will incorporate more nuanced psychology. Current models might include basic demographic and behavioral data, but future human Digital Twins may integrate psychological drivers, cognitive biases, and

social influences in more detail. Academic work (e.g., Muschkiet et al. 2022) is pushing in this direction, conceptualizing human Digital Twins not just as “virtual mirrors” of what people did, but as “virtual orchestrators” that simulate what people could do under various influences;

- **Standardization and Accessibility:** as the field matures, we can expect more standard tools and frameworks for creating digital twins in marketing, making it easier for non-tech companies to adopt;
- **Human Oversight and Ethical AI Focus:** in the future, there will be a stronger emphasis on ensuring these AI-driven twins are used responsibly. We foresee roles like “AI ethicist” or “simulation auditor” becoming part of the marketing analytics teams, to review how simulations are set up and how their results are applied, ensuring absence of bias and fairness.

Looking forward, it is evident that the integration of AI and Digital Twin architecture will increasingly redefine the methodological foundations by which organizations model and engage with consumer behavior. As future trends unfold – like hyper-personalized simulations and real-time decision intelligence – the organizations that build competence in this area stand to gain a significant competitive edge. Market research could become a living, continuous process, with AI agents constantly exploring the “what-ifs” of business on our behalf. In such a world, **human decision-makers will be better informed than ever, but their role will also evolve**, focusing on creativity, ethics, and strategic judgment, while letting AI handle the heavy analytical lifting.

## Disclosure

The authors are TP Infinity employees, and TWIN methodology has been developed as a company tool.

## Acknowledgements

The authors would like to express their sincere gratitude to **TP Infinity** for providing essential support and access to the tools and resources necessary for the development of this research. We also wish to thank all individuals who directly and indirectly contributed to this work through their insights, collaboration, or assistance. Finally, we extend our appreciation to the Applied Statistics Association for giving us the valuable opportunity to present our findings to their esteemed members.

## References

- BARRICELLI, B. R., CASIRAGHI, E., & FOGLI, D. (2019). A survey on digital twin: Definitions, characteristics, applications, and design implications. *IEEE Access*, 7, 167653-167671.
- SPERLING, J. (2023). AI-Generated Digital Twins: Shaping the Future of Business – Summary of a CBS workshop. Retrieved from <https://business.columbia.edu/insights/digital-future/ai-agent-digital-twin#:~:text=Peng%20explained%20that%20when%20simulating,traits%20with%20specific%20product%20types>
- MORAN, J. (2022). Digital Twins: A Marketer’s Guide. Retrieved from [https://www.destination-crm.com/Articles/ReadArticle.aspx?ArticleID=155818#.Y2qc1wPhh\\_8.linkedin](https://www.destination-crm.com/Articles/ReadArticle.aspx?ArticleID=155818#.Y2qc1wPhh_8.linkedin)
- BOURNE, J. (2025). Digital twins evolve to enable hyper-personalized advertising. Retrieved from <https://www.emarketer.com/content/2025-trend-digital-twins-enable-hyper-personalized-advertising>
- DUERST, M., COLBORNE, C., & VELOSA, A. (2022). A digital twin of a customer predicts the best consumer experience. *Gartner Business Quarterly*. Retrieved from: <https://www.gartner.com/en/insights/gartner-business-quarterly/q2-2022/digital-twin-of-a-customer>
- KORST, J., PUNTONI, S., & TOUBIA, O. (2025). How Gen AI is transforming market research. *Harvard Business Review*, 103(3), 90-99.
- MUSCHKIET, M., PASCHMANN, J. W., & NISSEN, A. (2022). Towards human digital twins for improving customer experience. *ICIS Proceedings*. 4. Retrieved from [https://aisel.aisnet.org/icis2022/serv\\_science/serv\\_science/4](https://aisel.aisnet.org/icis2022/serv_science/serv_science/4)
- TTEC DIGITAL (2023). How a digital twin of a customer can inform stronger customer experiences in retail. Retrieved from <https://www.ttecdigital.com/articles/how-a-digital-twin-of-a-customer-can-inform-stronger-customer-experiences-in-retail#:~:text=,%E2%80%9D>

## **COUNTERACTING FUTURE CLIMATE CHALLENGES USING ARTIFICIAL INTELLIGENCE MODELS: AN IRISH CASE STUDY**

### **CONTRASTARE LE SFIDE CLIMATICHE FUTURE UTILIZZANDO MODELLI DI INTELLIGENZA ARTIFICIALE: UN CASO DI STUDIO IRLANDESE**

*Yuri Calleo<sup>1</sup>, Francesco Pilla<sup>2</sup>, Simone Di Zio<sup>3</sup>*

#### **Sommario**

Questo studio esplora un approccio innovativo per affrontare le sfide climatiche future attraverso l'integrazione di metodi di decisione spaziale e modelli di intelligenza artificiale. Attraverso l'utilizzo del Real-Time Spatial Delphi (RTSD), il lavoro propone una metodologia ibrida che combina la mappatura collaborativa in tempo reale con strumenti testo-immagine e modelli generativi pre-addestrati, per visualizzare scenari futuri e generare proposte di policy. Lo studio è stato applicato ad un caso irlandese, coinvolgendo esperti delle città di Cork, Galway e Limerick per identificare aree vulnerabili a rischi climatici entro il 2050. I risultati mostrano elevati livelli di consenso spaziale e producono visualizzazioni realistiche dei futuri scenari ambientali. Le implicazioni principali riguardano la maggiore comprensibilità dei risultati per i decisori pubblici e una più rapida elaborazione di strategie di adattamento. Il metodo proposto colma un vuoto nella letteratura sul Foresight strategico, contribuendo alla democratizzazione della conoscenza spaziale e alla definizione di politiche resilienti.

#### **Abstract**

*This study explores an innovative approach to tackling future climate challenges by integrating spatial decision-making methods and artificial intelligence models. Building upon the Real-Time Spatial Delphi (RTSD) method, we propose a hybrid methodology combining real-time collaborative mapping with text-to-image tools and generative models to visualize future scenarios and draft policy proposals. The method was validated through a case study in Ireland involving experts from Cork, Galway, and Limerick, who identified vulnerable areas to climate risks by 2050. Results indicate a high level of spatial consensus and provide realistic visualizations of*

<sup>1</sup> University "G. d'Annunzio", Department of Economics, Chieti-Pescara, Italy - e-mail: yuri.calleo@unich.it

<sup>2</sup> University College Dublin, School of Architecture, Planning and Environmental Policy, Ireland - e-mail: francesco.pilla@ucd.ie

<sup>3</sup> University "G. d'Annunzio", DiSEGS, Chieti-Pescara, Italy - e-mail: s.dizio@unich.it

*projected environmental threats. The main implications involve improved communication of results to policymakers and faster development of adaptation strategies. This hybrid approach addresses key gaps in the strategic foresight literature and enhances the usability of spatial knowledge for effective, resilient policy design.*

## 1. Introduzione

Il cambiamento climatico presenta delle sfide complesse e articolate che influenzano in misura crescente i sistemi ecologici, sociali ed economici su scala globale. Questi cambiamenti minacciano costantemente la biodiversità e aggravano l'insicurezza idrica e alimentare, comportando seri rischi per la salute. Inoltre, le popolazioni vulnerabili, in particolare quelle a basso reddito o delle regioni costiere, sono colpite in modo sproporzionato a causa della loro limitata capacità di adattamento (Toimil *et al.*, 2020). Affrontare queste sfide richiede una comprensione approfondita delle dinamiche climatiche, strategie efficaci di mitigazione per ridurre le emissioni e quadri adattivi in grado di rafforzare la resilienza in tutti i settori. In questo campo, la decisione spaziale (in inglese SDM, Spatial Decision-Making) si riferisce al processo di scelta tra diverse alternative – comuni o non – tenendo conto dei modelli spaziali, delle informazioni geografiche e della distribuzione di risorse nello spazio. Questo approccio richiede una consapevolezza dei fattori spaziali, come la localizzazione, la prossimità e le interrelazioni tra i diversi elementi. Come affermato da Papadimitriou (2020), poiché la SDM implica spesso interazioni complesse ed imprevedibili all'interno di un contesto geografico specifico, è fondamentale considerare la disposizione spaziale delle variabili al suo interno che molto spesso comprende oggetti fisici, organismi viventi e processi dinamici. Di conseguenza, la complessità spaziale deriva da una molteplicità di fattori, sia naturali che sociali, elementi essenziali per l'analisi di potenziali scenari distopici e per orientare consapevolmente le azioni nel presente (Kohler, 2021).

Nel contesto degli studi sui futuri (in inglese FS, Futures Studies), gli scenari spaziali vengono spesso elaborati per rappresentare potenziali configurazioni territoriali in condizioni complesse, utilizzando modelli quantitativi, simulazioni o strumenti analitici con l'obiettivo di prevedere e/o anticipare i cambiamenti dinamici nel tempo (es. Calleo *et al.*, 2023). Tuttavia, data l'imprevedibilità del futuro, anche i modelli statistici più avanzati non garantiscono previsioni affidabili nel lungo periodo. Per questo motivo, nei FS si è passati da approcci previsionali tradizionali a pratiche di foresight strategico (in inglese, meglio definito come Strategic Foresight), volte ad esplorare una pluralità di futuri possibili ed a sviluppare capacità di anticipazione data la malleabilità degli scenari (vedi Kosow e Gaßner, 2008). Tra i diversi metodi

utilizzati per sviluppare scenari spaziali, il Real-Time Spatial Delphi (RTSD) (Di Zio *et al.*, 2017) si distingue come evoluzione del metodo Delphi tradizionale, adattato a contesti territoriali complessi. Attualmente, l'RTSD viene applicato in ambienti spaziali virtuali (Calleo e Pilla, 2024), nei quali gruppi di esperti interagiscono in tempo reale, rispondendo a specifici quesiti mediante l'inserimento di uno o più punti su una mappa dinamica. Questo sistema favorisce la collaborazione sincrona e la costruzione di un consenso spaziale tra gli esperti coinvolti, misurato tramite un elemento geometrico rappresentato da un cerchio  $C$ , che corrisponde all'indicatore più comunemente utilizzato nel metodo Delphi, ovvero l'intervallo interquartile ( $IQR$ ).

In genere, gli esperti collaborano per raggiungere un consenso spaziale, individuando un'area all'interno di un territorio specifico. Il processo, in seguito, genera dei principali output: a) risultati spaziali (es. mappe geografiche, mappe di calore, etc.); b) risultati derivanti da indicatori statistici (es. misure del livello di consenso, valori degli indicatori, etc.); c) commenti degli esperti (es. commenti testuali che supportano i punti spaziali). Benché il RTSD sia stato ampiamente utilizzato in letteratura, data l'imprevedibilità del futuro, anche solo immaginarlo risulta complesso, ed una delle maggiori complicanze del metodo è la comunicazione e diffusione dei risultati principali ai non esperti (es. autorità locali o cittadini). Difatti, un pubblico non esperto potrebbe non comprendere appieno il significato dei risultati statistici o derivanti dai quesiti spaziali, risultando in una minor consapevolezza delle problematiche espone. Inoltre, come evidenziato dalla letteratura scientifica, l'RTSD non dispone di una metodologia strutturata per l'elaborazione di politiche da prendere in considerazione successivamente al processo di decisione spaziale.

Per affrontare questa sfida proponiamo un'evoluzione del metodo RTSD introducendo un approccio ibrido innovativo che integra il processo standard con strumenti basati sull'intelligenza artificiale. Il metodo proposto utilizza modelli testo-immagine per la generazione di rappresentazioni visive – sotto forma di immagini – di possibili scenari futuri e modelli generativi pre-addestrati per la formulazione di una bozza di politiche di contrasto. Nello specifico gli obiettivi di ricerca sono:

- Migliorare il Real-Time Spatial Delphi attraverso l'integrazione di modelli di Intelligenza
- Artificiale nelle fasi conclusive del processo.
- Utilizzare modelli di generazione testo-immagine per creare rappresentazioni visive realistiche di scenari, basandosi sui risultati ottenuti dal metodo Real-Time Spatial Delphi.
- Considerare gli output del Real-Time Spatial Delphi utilizzando i trasformatori generativi pre-addestrati, al fine di produrre una bozza preliminare di politiche territoriali da sottoporre a valutazione.

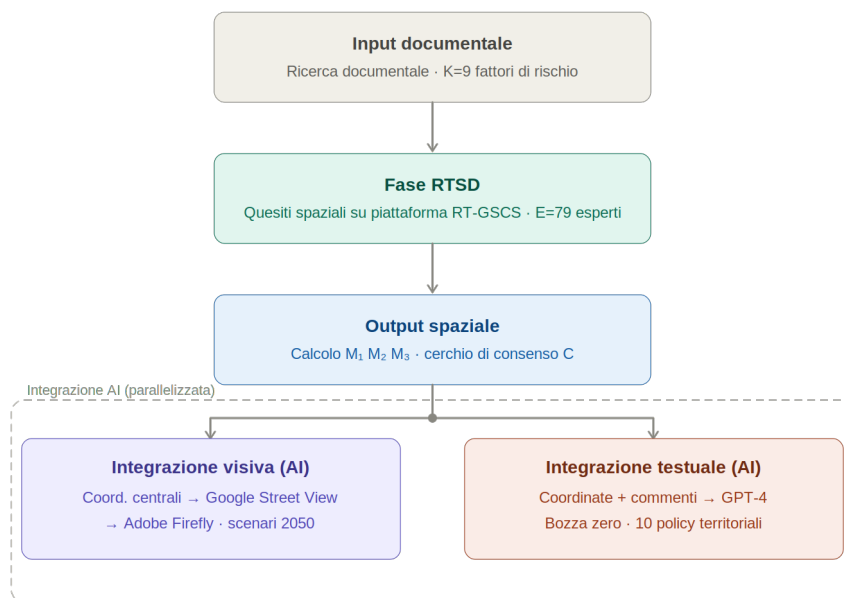
L'obiettivo di tale integrazione è quello di rendere più efficiente la costruzione degli scenari spaziali, migliorandone la comprensibilità e l'applicabilità per un'implementazione tempestiva da parte dei decisori politici. Rispetto ai tradizionali metodi di spatial decision-making o foresight, che si limitano a produrre mappe di rischio o scenari statistici, il nostro approccio ibrido colma la lacuna operativa tra la visione spaziale e l'azione politica, strutturando un framework dove l'output quantitativo diventa immediatamente l'input qualitativo per l'IA. In questo contributo, presentiamo uno studio di caso incentrato sulle minacce legate al cambiamento climatico nelle città di Cork, Galway e Limerick, in Irlanda, con proiezione al 2050.

## 2. Metodologia

### *Definizione del metodo*

In questo studio proponiamo un metodo ibrido innovativo partendo dalla considerazione dei rischi climatici attualmente in atto, seguita da un'approfondita ricerca documentale finalizzata alla raccolta di letteratura rilevante ed all'elaborazione di analisi spaziali (Fig. 1). Il metodo viene validato attraverso un caso di studio condotto in Irlanda, coinvolgendo esperti del settore e competenti nelle città di Cork, Galway e Limerick, con l'obiettivo di affrontare sfide regionali diversificate e contribuire alla definizione di politiche climatiche specifiche sia a livello locale che nazionale per l'anno 2050 (in linea con gli studi sul cambiamento climatico delle Nazioni Unite e dell'Unione Europea). La ricerca documentale ha inizio con una revisione accurata della letteratura scientifica, dei report europei, consultazione con stakeholder e lettura dei dati rilevanti provenienti da statistiche ufficiali (es. Central Statistics Office, <https://www.cso.ie/en/index.html>). Da quanto emerge, le criticità sono differenti tra le tre città, dove si hanno problematiche relative ad alluvioni, nevicate, tempeste di vento improvvise ed innalzamento del livello del mare, eventi che possono avere un impatto significativo sull'ambiente e sulla popolazione. Questi fattori costituiscono la base per la formulazione delle domande di ricerca da porre al nostro panel di esperti all'interno della piattaforma RTSD. Il sistema adottato in questo studio è il Real-Time Geo-Spatial Consensus System (RT-GSCS – [www.rtgscs.com](http://www.rtgscs.com)), una piattaforma web open source che consente la somministrazione in tempo reale di questionari spaziali, seguendo la logica del metodo RTSD (Calleo *et al.*, 2023).

Figura 1. Diagramma del metodo ibrido RTSD-AI



Una volta definito il quadro metodologico e raccolta la letteratura, i dati spaziali e consultati gli esperti, si procede con la compilazione di una lista preliminare di fattori ambientali chiave ( $K$ ) considerati come potenziali minacce per le città analizzate. Tali fattori sono implicitamente inclusi nei piani territoriali approvati a livello municipale, facilmente reperibili attraverso report online aggiornati. Per ciascuna città, sono state identificate le tre minacce principali con priorità più alta, risultando in un totale di  $K = 9$  fattori di rischio ambientale. Di conseguenza sono state formulate  $RQ = 9$  domande di ricerca da sottoporre al panel di esperti attraverso la piattaforma RT-GSCS e corredate da coordinate spaziali  $(x, y)$  all'interno della stessa. Una panoramica delle domande proposte è illustrata nella Tab. 1.

Una volta caricate correttamente le domande nel sistema, si procede all'individuazione del panel di esperti. Le convocazioni vengono inviate tramite e-mail, contenenti dettagli tecnici completi, scadenze, linee guida e istruzioni per l'utilizzo della piattaforma. Il panel è composto da 176 esperti invitati, dei quali  $E = 79$  hanno effettivamente partecipato, contribuendo almeno con l'inserimento di  $n_i$  punti spaziali. Il gruppo include esperti accademici, rappresentanti delle autorità locali, membri del settore industriale, associativi ambientali e specialisti in foresight strategico. Tra i partecipanti, 65 sono di genere femminile, affiliati o provenienti dalla città di riferimento, garantendo così un panel eterogeneo e rappresentativo rispetto agli obiettivi della ricerca. Il sondaggio inizia utilizzando le e-mail degli esperti registrate nel sistema, che permettono di accedere

al link della sessione privata fornito dai ricercatori. Una volta entrati nel sistema, gli esperti si trovano di fronte ad un questionario spaziale in tempo reale, caratterizzato da una mappa dinamica e da strumenti di supporto (es. informazioni dei risultati in tempo reale, possibilità di cambiare la visualizzazione della mappa, strumenti di inserimento delle risposte ecc.). La barra laterale inoltre evidenzia gli elementi fondamentali, come le valutazioni comunali, studi precedenti, analisi spaziali di riferimento e le linee guida del sondaggio. Gli esperti possono quindi rispondere alle domande in base alla propria area di competenze e, selezionata una domanda, hanno la possibilità di collocare uno o più punti  $z_i$  sulla mappa. Tutti i contributi sono aperti a commenti anonimi in tempo reale, favorendo così un confronto dinamico tra i partecipanti.

Tabella 1. Domande proposte (tradotte in lingua italiana)

Area	Domande di ricerca		
	$RQ_1$	$RQ_2$	$RQ_3$
Cork	Quale area sarà più a rischio di inondazioni?	Quale area sarà più a rischio per l'impatto delle nevicate?	Quale area sarà più a rischio per l'impatto delle tempeste di vento?
Galway	Quale area sarà più a rischio di inondazioni?	Quale area sarà più a rischio per l'impatto delle nevicate?	Quale area sarà più a rischio per l'impatto delle tempeste di vento?
Limerick	Quale area sarà più a rischio di inondazioni?	Quale area sarà più a rischio per l'impatto dell'innalzamento del livello del mare?	Quale area sarà più a rischio per l'impatto delle precipitazioni?

### Algoritmi del Real-Time Spatial Delphi e modelli di intelligenza artificiale

L'algoritmo implementato nella piattaforma RT-GSCS segue la logica consigliata da Di Zio et al. (2017), ovvero quando un punto  $z_i$  viene posizionato sulla mappa dinamica, gli esperti osservano un cerchio dinamico  $C_i$  che si adatta in dimensione e posizione in base agli altri input presenti. Come nella metodologia Delphi tradizionale, se un punto  $z_i$  viene collocato al di fuori del cerchio  $C_i$ , è richiesta una giustificazione testuale; tali commenti sono evidenziati con una bandierina rossa per segnalare la divergenza e stimolare il raggiungimento del consenso (ove possibile e/o necessario). Nel nostro caso, poiché il processo avviene in tempo reale, il raggiungimento del consenso non può essere considerato l'unico criterio di conclusione del processo decisionale, ma è altrettanto importante garantire la stabilità. Per questo motivo, viene assegnato agli esperti un periodo lavorativo di 30 giorni, con i primi 20 giorni dedicati alla partecipazione attiva e gli ultimi 10 focalizzati sul raggiungimento della stabilità. Durante questa fase finale, gli esperti rientrano sulla piattaforma, analizzano attentamente i pattern di risposta e

affinano, ove congruente, le aree di consenso già formate o forniscono maggiori informazioni a riguardo.

Dopo aver acquisito i risultati spaziali e non spaziali, si procede con l'integrazione di modelli di intelligenza artificiale nel processo, al fine di visualizzare le decisioni in forma "reale" e formulare una bozza di politiche. Per ottenere una rappresentazione visiva, tramite immagini, degli scenari spaziali delineati, si considera il cerchio finale  $C$ , derivante dal processo per ogni domanda di ricerca  $RQ$ , identificando un'area minima

$\min(A)$ , dove  $A = [A_1, A_2, \dots, A_N]$  è il vettore delle aree relative alla sequenza di cerchi di consenso. Le aree emerse rappresentano la localizzazione specifica indicata dagli esperti come rilevante per concentrare l'attenzione politica sulle minacce considerate. Considerando la distribuzione di  $Z = Z_1, Z_2, \dots, Z_N$  punti con  $x, y$  coordinate associate, per validare il nostro metodo estraiamo le coordinate del punto  $z_i$  centrale nel cerchio finale  $C$ . Queste coordinate rappresentano un'area specifica del territorio e vengono utilizzate per estrarre tramite Google Street View, l'immagine reale del luogo (es. una strada, un quartiere o un'area specifica). Per ciascuna domanda, una volta ottenuta l'immagine  $I$ , procediamo ad importarla in Adobe Firefly Image ([www.firefly.adobe.com](http://www.firefly.adobe.com)), utilizzando lo strumento "Generative Fill", e fornendo un prompt per generare una nuova immagine che rappresenti uno scenario plausibile. Per ridurre al minimo i possibili bias legati alla scelta del prompt, adottiamo i seguenti prompt:

- Cork: "Genera un'area allagata" ( $RQ_1$ ), "Genera un'area colpita da nevicate" ( $RQ_2$ ), "Genera un'area interessata da tempeste di vento" ( $RQ_3$ ).
- Galway: "Genera un'area allagata" ( $RQ_1$ ), "Genera un'area colpita da nevicate" ( $RQ_2$ ), "Genera un'area interessata da tempeste di vento" ( $RQ_3$ ).
- Limerick: "Genera un'area allagata" ( $RQ_1$ ), "Genera un'area interessata dall'innalzamento del livello del mare" ( $RQ_2$ ), "Genera un'area soggetta a forti precipitazioni" ( $RQ_3$ ).

In questo modo, il modello produce tre distinti scenari visivi futuri per ogni domanda, pronti per essere comunicati e diffusi. Tuttavia, negli studi sui futuri, lo sviluppo di scenari dovrebbe enfatizzare la formulazione attiva di politiche, con l'obiettivo di contrastare o mitigare le minacce future e implementare strategie presenti. Per superare questa limitazione all'interno del RTSD utilizziamo GPT-4 ([www.openai.com](http://www.openai.com)) per generare una lista preliminare di politiche basata sugli scenari individuati dagli output del RTSD. Il prompt fornito al GPT è strutturato considerando i risultati ottenuti dal processo decisionale ed in particolare: (1) Fattori chiave: identificazione del rischio climatico considerato; (2) Coordinate: informazioni sull'area individuata dal punto centrale  $z_c$  corrispondente all'area  $\min(A)$ . Le coordinate fungono da ponte algoritmico tra l'o-

pinione spaziale dell'esperto e l'ancoraggio alla realtà visiva, permettendo al modello generativo di non inventare un paesaggio generico, ma di alterare l'effettiva morfologia del territorio selezionato; (3) Commenti: tutti i feedback in relazione alla risposta spaziale fornita dagli esperti (3). A questo punto, la descrizione del prompt riveste un ruolo cruciale nel determinare gli esiti prodotti dal GPT, e di conseguenza adottiamo la seguente formulazione:

- **P1:** “*Agisci come un decisore politico... genera una lista di 10 politiche volte a mitigare gli impatti previsti per il 2050 sulla base delle seguenti informazioni... (1), (2), (3)*”.

La scelta di estrarre solamente 10 politiche è meramente soggettiva e rappresenta un tentativo di validazione ed esplicazione del metodo da utilizzare come base di discussione tra i policymakers, e non come politiche definitive da implementare. Difatti, questo approccio colma una lacuna significativa nella letteratura scientifica riguardante le fasi finali del processo RTSD, dove i risultati vengono presentati e le politiche formulate. Grazie a questo metodo, anche persone senza competenze specialistiche possono acquisire una visione dei possibili scenari futuri, favorendo una consapevolezza più ampia. Inoltre, consente di ridurre tempi e risorse generalmente necessari per workshop o round iterativi finalizzati a generare una bozza di politiche territoriali. Questo approccio strutturato migliora l'efficienza e l'accessibilità nella formulazione delle politiche, garantendo una risposta solida alle sfide ambientali previste.

### 3. Risultati principali

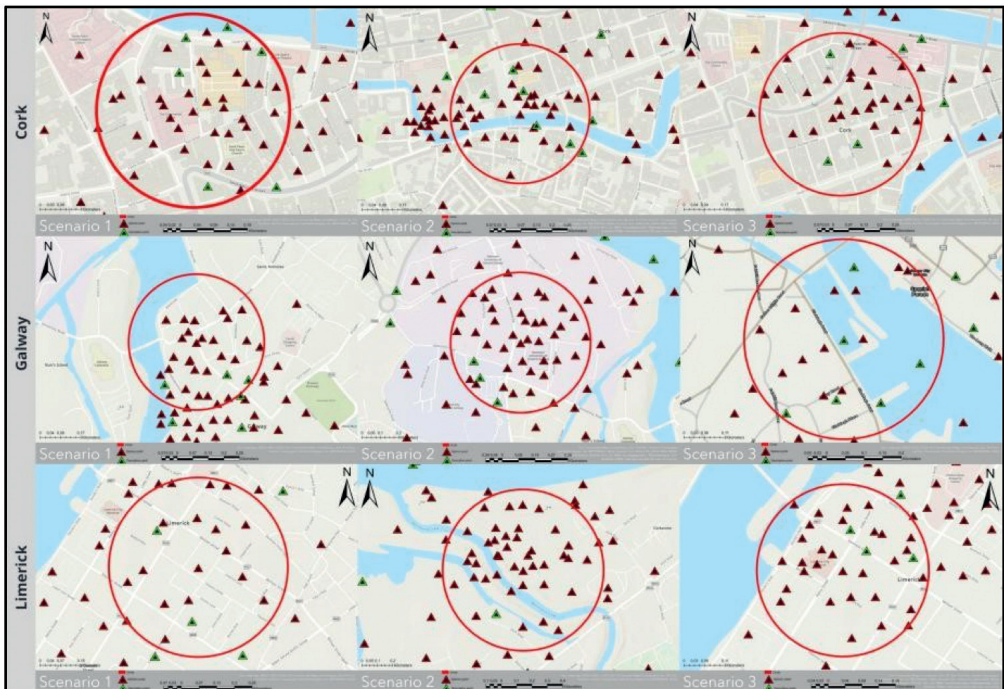
#### *Analisi spaziali risultanti dal processo decisionale*

La sessione panel è iniziata il 1° febbraio 2024 e si è conclusa il 2 marzo 2024, coprendo un totale di  $T = 30$  giorni, durante i quali gli esperti hanno raggiunto un significativo livello di consenso e hanno risposto a tutte le domande presenti nel sistema con oltre  $Z = 604$  punti e 96 commenti. I risultati ottenuti dalle analisi spaziali sono raffigurati nella Fig. 2, e rappresentano la collaborazione degli esperti nell'identificazione delle aree di maggior rischio ambientale e climatico alla fine del procedimento.

Per Cork, gli esperti hanno identificato tre scenari potenziali che evidenziano tali minacce previste entro il 2050, comprendendo un rischio di inondazione nei pressi del Fiume Lee (Scenario 1), vulnerabilità legate ad improvvise nevicate e tempeste nel centro città (Scenario 2 e Scenario 3). Dai commenti, tali scenari sono pericolosi per la popolazione a causa di danni strutturali, rischi legati alla caduta di alberi ed impatti sui trasporti e sulla sicurezza pubblica. Per la città di Galway, gli esperti hanno identificato tre scenari: una vulnerabilità agli eventi di allagamento nelle vicinanze del fiume Corrib (Scenario 1), rischi derivanti da nevicate improvvise nel centro città (Scenario

2) e preoccupazioni legate a intense tempeste di vento nell'area sudorientale di Galway (Scenario 3). Infine, per Limerick, gli esperti hanno espresso preoccupazioni riguardo ai rischi di allagamento vicino al fiume Shannon (Scenario 1), alle conseguenze dell'innalzamento del livello del mare presso il punto di accesso alla contea e agli impatti potenziali di eventi di pioggia improvvisa nel centro città. Complessivamente, le informazioni spaziali fornite dagli esperti evidenziano possibili minacce climatiche per il 2050, pronte per essere valutate da policymaker ed enti governativi. Tuttavia, essi non possono essere considerati l'unico output da valutare, ma ulteriori aspetti devono essere illustrati, incluse le metriche di consenso e le successive implementazioni tramite il supporto di modelli di intelligenza artificiale. In questo caso, sono stati considerati tre indicatori principali  $M_1(FCkm^2)$  dove  $FC$  rappresenta l'area del cerchio finale;  $M_2 = 1 - FC/S$ , calcolato come il rapporto tra l'area del cerchio finale  $FC$  e la superficie della città  $S$ .

Figura 2. Risultati spaziali derivanti dal processo di convergenza



Un valore dell'indicatore  $M_2 \rightarrow 1$  suggerisce che più piccolo sarà  $FC$  in comparazione ad  $S$ , indicando quindi un consenso spaziale. L'ultimo indicatore  $M_3 (FC/IC \cdot 100)$ , misura il grado di consenso tra gli esperti ed è espresso in percentuale, calcolando il

rapporto tra l'area del cerchio finale  $FC$  e l'area del cerchio iniziale  $IC$ . In questo indicatore, valori più alti (vicini al 100%) indicano una bassa convergenza delle opinioni tra i partecipanti, mentre valori prossimi a zero indicano un alto livello di consenso raggiunto. Da un punto di vista operativo per i policymaker, un alto livello di consenso ( $M_3 \rightarrow 0$ ) combinato a una piccola area finale ( $M_1$  e  $M_2$ ) identifica una zona rossa ad altissima priorità, permettendo di allocare fondi e risorse in modo chirurgico e non dispersivo.

Tabella 2. Misure di consenso spaziale

Città	Scenario	$S$ (km <sup>2</sup> )	$IC$	$FC$ ( $M_1$ )	$M_2$	$M_3$ (%)	$Z$	$M$
Cork	Sc.1	187	1.41	0.16	0.999	11%	71	10
Cork	Sc.2	187	2.28	0.18	0.999	7%	88	13
Cork	Sc.3	187	1.28	0.14	0.999	10%	58	9
Galway	Sc.1	57.3	1.32	0.17	0.997	12%	80	8
Galway	Sc.2	57.3	2.42	0.10	0.998	4%	90	12
Galway	Sc.3	57.3	1.51	0.16	0.997	10%	58	9
Limerick	Sc.1	61.3	1.41	0.20	0.996	14%	44	6
Limerick	Sc.2	61.3	3.10	0.32	0.994	10%	60	8
Limerick	Sc.3	61.3	2.55	0.14	0.997	5%	55	11
<b>Total</b>							<b>604</b>	<b>96</b>

Nella Tab. 2, è possibile osservare un forte consenso, basato sulla riduzione dell'area per tutti e tre gli scenari confrontando  $IC$  con l'indicatore  $M_1$ . Inoltre, per tutte le domande – e di conseguenza gli scenari –  $M_2 \rightarrow 1$  suggerendo un forte consenso spaziale. Infine, per i valori di  $M_3$  abbiamo dei valori molto bassi suggerendo un consenso tra gli esperti molto alto (nella letteratura scientifica possiamo definire il consenso quando  $IQR < 20\%$ ).

### Output provenienti dai modelli di Intelligenza Artificiale

A seguito della presentazione dei risultati principali ottenuti dalle analisi precedenti, abbiamo utilizzato modelli testo-immagine per la visualizzazione di scenari futuri potenziali, con un tempo medio di generazione pari a  $\bar{x} = 30$  secondi. La Fig. 3 illustra nove scenari distinti per ciascuna città, integrando immagini territoriali autentiche arricchite con elementi generati dall'intelligenza artificiale.

Dalla Fig. 3, è possibile constatare come lo Scenario 1 in tutte e tre le città simula efficacemente eventi di possibili inondazioni, offrendo una rappresentazione realistica delle condizioni previste per l'anno 2050. Vi è da sottolineare che queste immagini evitano proiezioni utopiche, ma forniscono una rappresentazione chiara e accessibile

dei rischi imminenti. Per quanto riguarda gli impatti atmosferici, i modelli raffigurano diverse sfide future: lo scenario 2 per Cork e Galway mostra potenziali nevicate, lo scenario 3 per Cork e Galway illustra possibili tempeste di vento, mentre lo Scenario 3 per Limerick rappresenta gli effetti di forti piogge. Infine, lo Scenario 2 per Limerick visualizza l'innalzamento previsto del livello del mare, con conseguenti rischi di allagamento lungo gli argini dei fiumi. Queste rappresentazioni visive contribuiscono efficacemente a comunicare le sfide ambientali potenziali che le città e i loro abitanti potrebbero affrontare in futuro. Questa implementazione rappresenta un'innovazione significativa all'interno del metodo RTSD, estendendo il suo ambito di applicazione oltre la pianificazione territoriale a settori diversi come lo sviluppo di prodotti, le strategie aziendali e la pianificazione della sicurezza. I tradizionali software di graphic design spesso sono soggetti a bias introdotti dai progettisti umani e a inefficienze nell'uso delle risorse. Al contrario, i modelli di intelligenza artificiale – pur essendo ancora in fase di sviluppo e con alcune limitazioni – possiedono un potenziale significativo per migliorare i processi decisionali.

Dopo la fase di visioning, integriamo i risultati in GPT-4 per generare una bozza di politiche potenziali, da perfezionare e discutere con i decisori politici e le autorità locali. Un esempio delle politiche sviluppate è illustrato nella Fig. 4 (per i risultati completi si rimanda a questo link: <https://tinyurl.com/nrjx47ap>).

Da quanto emerso, secondo il modello di intelligenza artificiale, le strategie si concentrano su diversi approcci chiave: per quanto concerne il rischio di allagamenti vi sono suggerimenti in investimenti in difesa contro le inondazioni, adozione di soluzioni verdi, coinvolgimento della comunità e pianificazione urbanistica consapevole del clima per rafforzare la resilienza urbana. Analogamente, per prepararsi a potenziali tempeste di neve, le politiche enfatizzano la preparazione invernale, la resilienza delle infrastrutture e la partecipazione comunitaria per garantire la sicurezza pubblica e la continuità dei servizi. Per le vulnerabilità legate alle tempeste di vento, le strategie includono l'applicazione di norme edilizie resilienti, il miglioramento della prontezza alle emergenze, l'investimento in infrastrutture verdi e la promozione della collaborazione intersettoriale per ridurre i danni e costruire la resilienza della comunità.

Figura 3. Scenari visuali ottenuti da modelli testo-immagine

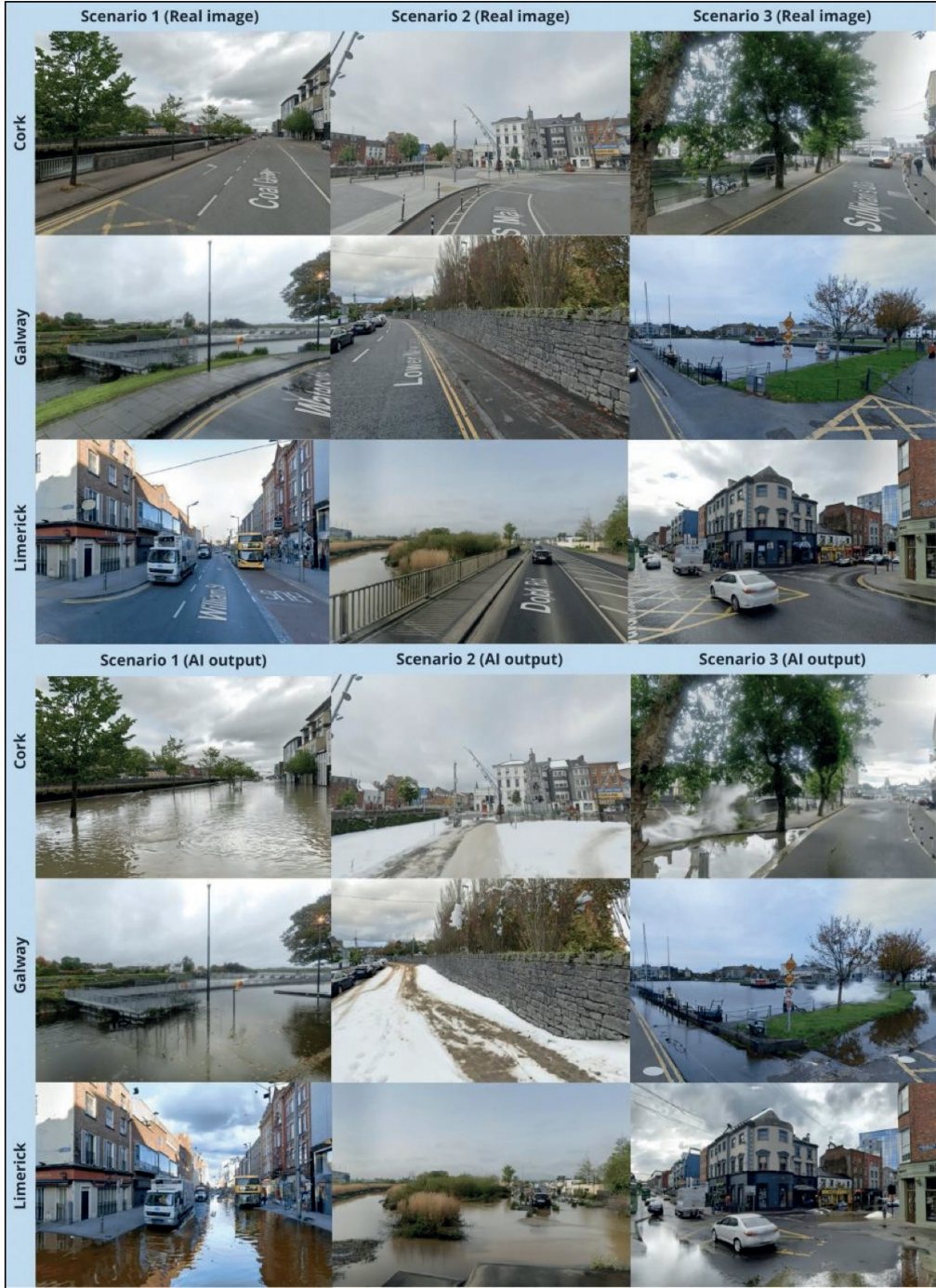


Figura 4. Bozza delle politiche emerse da Chat-GPT 4 (Sc. I Cork)

<ol style="list-style-type: none"> <li>1. <b>Investimenti nelle Infrastrutture di Difesa contro le Alluvioni:</b> Assegnare fondi per la costruzione e la manutenzione di sistemi robusti di difesa contro le alluvioni, come barriere, argini e infrastrutture per la gestione delle acque meteoriche, al fine di proteggere le aree vulnerabili di Cork dalle inondazioni.</li> <li>2. <b>Sviluppo di Infrastrutture Verdi:</b> Promuovere l'implementazione di soluzioni di infrastruttura verde, come tetti verdi, pavimentazioni permeabili e spazi verdi urbani, per assorbire le precipitazioni, ridurre il deflusso e mitigare i rischi di alluvione, migliorando al contempo la biodiversità e l'estetica urbana.</li> <li>3. <b>Gestione delle Pianure Alluvionali e Pianificazione Urbanistica:</b> Attuare regolamenti rigorosi sull'uso del suolo e lo sviluppo nelle aree soggette a inondazioni, inclusi vincoli urbanistici e incentivi per pratiche edilizie resilienti, al fine di ridurre l'esposizione ai rischi di alluvione.</li> <li>4. <b>Normative Edilizie Resilienti al Clima:</b> Applicare codici edilizi severi che richiedano alle nuove costruzioni e ristrutturazioni di includere caratteristiche progettuali resilienti al clima, come fondazioni sopraelevate, misure di protezione contro le alluvioni e materiali resistenti, per fronteggiare eventi alluvionali.</li> <li>5. <b>Campagne di Sensibilizzazione e Educazione Pubblica:</b> Avviare campagne di sensibilizzazione complete per educare i residenti sui rischi di alluvione, la preparazione alle emergenze, i percorsi di evacuazione e l'importanza di adottare comportamenti e misure resilienti alle alluvioni a livello individuale e comunitario.</li> </ol>	<ol style="list-style-type: none"> <li>6. <b>Integrazione delle Considerazioni Climatiche nella Pianificazione Urbana:</b> Integrare le considerazioni legate al cambiamento climatico nei processi di pianificazione urbana, dando priorità all'adattamento climatico e alla resilienza nei piani di sviluppo, nelle infrastrutture di trasporto e nelle decisioni sull'uso del territorio, per ridurre l'esposizione alle inondazioni e rafforzare la resilienza della città.</li> <li>7. <b>Soluzioni Basate sulla Natura:</b> Attuare soluzioni basate sulla natura come il ripristino delle zone umide, la creazione di fasce vegetative lungo i corsi d'acqua e il potenziamento della capacità naturale delle pianure alluvionali di immagazzinare l'acqua, per attenuare le alluvioni, migliorare la qualità dell'acqua e fornire benefici ecologici aggiuntivi, riducendo al contempo i rischi di inondazione a Cork.</li> <li>8. <b>Investimenti in Sistemi di Allerta Precoce:</b> Investire in tecnologie avanzate e sistemi di allerta precoce per la previsione e il monitoraggio delle alluvioni, al fine di fornire avvisi tempestivi a residenti, soccorritori e autorità locali, consentendo risposte proattive e riducendo danni e perdite dovuti alle inondazioni.</li> <li>9. <b>Coinvolgimento e Partecipazione della Comunità:</b> Favorire il coinvolgimento e la partecipazione della comunità attraverso iniziative di gestione del rischio alluvione basate sulla comunità, associazioni di quartiere e processi decisionali partecipativi, per responsabilizzare i residenti nell'identificare, dare priorità e attuare misure di resilienza adattate alle loro specifiche esigenze e circostanze.</li> <li>10. <b>Collaborazione Intersettoriale e Partenariati:</b> Stabilire partenariati collaborativi tra enti governativi, istituzioni accademiche, attori del settore privato e organizzazioni comunitarie per sfruttare risorse, competenze e innovazione nell'attuazione di strategie olistiche di gestione del rischio alluvionale, condividere buone pratiche e costruire una resilienza collettiva ai cambiamenti climatici.</li> </ol>
--	---

Queste politiche rappresentano sforzi proattivi per mitigare i rischi climatici e garantire la sostenibilità a lungo termine della città di fronte alle sfide ambientali in evoluzione. Nello specifico, è opportuno evidenziare che il processo di generazione delle linee guida tramite GPT-4 non si è basato su prompt generici o astratti, bensì su un rigoroso ancoraggio ai dati di contesto. L'input fornito al modello di intelligenza artificiale ha incluso, infatti, sia le coordinate geografiche esatte dei cluster di vulnerabilità emersi dalla fase RTSD, sia i commenti qualitativi puntuali espressi dagli esperti locali durante le sessioni di consultazione. Questo duplice vincolo informativo è risultato decisivo nel prevenire la produzione di raccomandazioni puramente teoriche o standardizzate, un rischio altrimenti frequente nell'uso di modelli linguistici di grande scala. Al contrario, tale ancoraggio territoriale ha permesso di declinare le 10 policy sulle criticità geomorfologiche e idrografiche uniche di ciascun sito d'indagine. Di conseguenza, le azioni proposte risultano modellate specificamente sulle dinamiche di esondazione e gestione del rischio nell'area del Fiume Lee per la città di Cork, sulle fragilità ambientali del bacino del Fiume Corrib per Galway, e sulle complesse problematiche idrologiche legate al corso del Fiume Shannon per Limerick, garantendo una reale aderenza alle specifiche esigenze di pianificazione e resilienza locale.

#### 4. Conclusioni

In questo studio abbiamo presentato un approccio innovativo per potenziare il metodo RTSD attraverso l'integrazione di modelli avanzati di Intelligenza Artificiale nelle fasi finali del processo. Tale integrazione intende affrontare le difficoltà legate alla comunicazione e diffusione efficace degli scenari spaziali, nonché alla formulazione di strategie e politiche operative. Grazie all'utilizzo combinato di modelli testo-immagine

e ChatGPT, viene colmato il divario tra gli aspetti testuali, spaziali e statistici, favorendo così una comprensione più completa e una visualizzazione più chiara degli scenari futuri. La metodologia proposta colma una lacuna importante nella pianificazione degli scenari, offrendo rappresentazioni visive realistiche delle configurazioni spaziali potenziali e producendo bozze di politiche territoriali da sottoporre a valutazione. Questo approccio consente una maggiore efficienza nel processo di elaborazione delle politiche, alleggerendo il carico decisionale dei responsabili politici e agevolando la definizione di misure più tempestive e consapevoli. Il nostro contributo apporta un valore significativo sia al metodo RTSD sia all'ambito più ampio degli studi di futuro. Uno degli elementi distintivi del nostro lavoro riguarda il miglioramento della componente visiva del processo. Attraverso i modelli testo-immagine, infatti, è possibile generare visualizzazioni credibili degli scenari futuri a partire dai dati spaziali forniti dagli esperti. Questo consente una comprensione più immediata e intuitiva delle possibili configurazioni territoriali, favorendo decisioni più informate. Tali strumenti di previsione si rivelano cruciali, poiché aiutano i decisori ad anticipare rischi potenziali in modo realistico e credibile, in linea con quanto avviene nelle "stanze operative delle politiche", dove la complessità degli scenari spesso rende difficile interpretare le sfide emergenti. In questo contesto, l'IA fornisce una base concreta e pertinente per il confronto, rendendo tangibili esperienze future attraverso simulazioni visive costruite con rigore.

Inoltre, l'integrazione di ChatGPT ci consente di generare elenchi preliminari di politiche territoriali potenziali, che possono essere successivamente valutati. Ciò accelera la fase di sviluppo delle politiche e semplifica il lavoro dei policy maker, offrendo loro spunti concreti provenienti da una pianificazione supportata dall'IA. La nostra proposta favorisce anche uno sviluppo più integrato degli scenari, unendo in modo coerente contenuti testuali, dati spaziali e componenti statistiche. Questo approccio olistico consente di formulare strategie più robuste e adattabili per affrontare sfide territoriali complesse. In sintesi, la nostra metodologia arricchisce il metodo RTSD grazie all'uso di tecnologie IA, fornendo agli attori coinvolti strumenti più approfonditi e operativi per interpretare le dinamiche spaziali future. In questo modo, non solo si rafforza la capacità di pianificazione degli scenari, ma si dà anche maggiore potere ai decisori nell'affrontare un futuro sempre più incerto e articolato, in linea con l'esigenza di sviluppare scenari a lungo termine in contesti che richiedono decisioni urgenti. Il metodo proposto è altamente scalabile e replicabile in altri contesti urbani. I prerequisiti fondamentali includono: la disponibilità di dati geospaziali di base per istruire la piattaforma; l'ingaggio di un panel eterogeneo di esperti locali per garantire la validità del consenso; e l'accesso ad API di modelli generativi standard (testo-immagine e LLM) per la fase di visioning e policy drafting. Pur rappresentando un passo avanti rilevante nell'ambito degli studi di futuro, il nostro lavoro apre anche nuove direzioni di ricerca, come: 1) l'ottimizzazione

dei modelli di IA: migliorare ulteriormente l'efficacia dei modelli testo-immagine e dei transformer generativi per ottenere immagini e suggerimenti politici ancora più realistici e accurati. 2) Il coinvolgimento degli stakeholder: includere il feedback di cittadini e decisori nel processo di sviluppo degli scenari per aumentare la pertinenza e l'applicabilità delle politiche generate. 3) La validazione su casi reali: testare e convalidare il metodo in contesti spaziali e scenari diversi per verificarne affidabilità ed efficacia concreta. 4) Le considerazioni etiche: affrontare temi legati alla trasparenza, ai bias e alla responsabilità nell'uso dell'IA nella pianificazione strategica. Ed infine, 5) la valutazione degli impatti a lungo termine: monitorare gli effetti delle politiche attuate sulla base degli scenari generati, al fine di affinare le strategie future. L'esplorazione di queste aree di ricerca potrà favorire una più ampia e responsabile adozione dell'IA nella pianificazione territoriale, contribuendo a strategie di sviluppo spaziale più sostenibili e resilienti. L'utilizzo di modelli generativi per la stesura delle policy non intende sostituire il decisore umano, ma piuttosto fornire una "bozza zero" (drafting) per accelerare i tavoli di concertazione. È fondamentale riconoscere i limiti intrinseci di tali strumenti, tra cui il rischio di bias ereditati dai dati di addestramento o la produzione di "allucinazioni" (suggerimenti plausibili ma inapplicabili localmente). Pertanto, l'IA agisce in questo quadro esclusivamente come strumento di supporto alle decisioni (*Decision Support System*), in cui la validazione finale, l'adattamento al contesto normativo e l'implementazione restano prerogative insindacabili dei policymaker.

## Bibliografia

- CALLEO, Y., & PILLA, F. (2024). Optimizing spatial survey administration adopting RT-GSCS: A statistical perspective on performance metrics. *MethodsX*, 12, 102578.
- CALLEO, Y., DI ZIO, S., & PILLA, F. (2023). Facilitating spatial consensus in complex future scenarios through Real-Time Spatial Delphi: A novel web-based open platform. *Futures & Foresight Science*, 5(3-4), e155.
- DI ZIO, S., ROSAS, J. D. C., & LAMELZA, L. (2017). Real Time Spatial Delphi: Fast convergence of experts' opinions on the territory. *Technological Forecasting and Social Change*, 115, 143-154.
- KOHLER, K. (2021). Strategic foresight: Knowledge, tools, and methods for the future. *CSS Risk and Resilience Reports*.
- KOSOW, H., & GASSNER, R. (2008). Methods of future and scenario analysis: overview, assessment, and selection criteria (Vol. 39, p. 133). *DEU*.
- PAPADIMITRIOU, F. (2020). Spatial complexity: Theory, mathematical methods and applications. *Springer Nature*.

TOIMIL, A., LOSADA, I. J., NICHOLLS, R. J., DALRYMPLE, R. A., & STIVE, M. J. (2020). Addressing the challenges of climate change risks and adaptation in coastal areas: A review. *Coastal Engineering*, 156, 103611.



PRINTED IN JUNE 2026  
ON BEHALF OF  
GIAPETO EDITORE

[www.giapeto.it](http://www.giapeto.it)