

11th Scientific Meeting of the SIS Group  
"Statistics for the Evaluation and Quality in Services"

# BOOK OF **SHORT PAPERS**

## Editors

Andrea Bucci

Alfredo Cartone

Adelia Evangelista

Andrea Marletta



**STATISTICAL METHODS  
FOR EVALUATION AND QUALITY:  
TECHNIQUES, TECHNOLOGIES AND TRENDS (T<sup>3</sup>)**

**IES 2023 - Statistical Methods for Evaluation and Quality:  
Techniques, Technologies and Trends (T<sup>3</sup>)**

## **BOOK OF SHORT PAPERS**

Editors: Andrea Bucci, Alfredo Cartone, Adelia Evangelista and Andrea Marletta

Book of Short papers  
11th International Conference **IES 2023**  
Statistical Methods for Evaluation and Quality: Techniques, Technologies and  
Trends (T<sup>3</sup>)

University 'G. d'Annunzio' of Chieti-Pescara



## Scientific Committee of the group of the Italian Statistical Society on Statistics for the Evaluation and Quality of Services - SVQS

Maurizio Carpita	University of Brescia
Pietro Amenta	University of Sannio
Matilde Bini	European University of Rome
Antonello D'Ambra	University of Campania 'L. Vanvitelli'
Luigi D'Ambra	University of Naples 'Federico II'
Paolo Mariani	University of Milan 'Bicocca'
Marica Manisera	University of Brescia
Monica Palma	University of Salento
Pasquale Sarnacchiaro	University of Naples 'Federico II'

## Program Committee of the conference IES 2023

<b>Chair:</b> Paolo Mariani	University of Milan 'Bicocca'
Matilde Bini	European University of Rome
Mario Bolzan	University of Padova
Giuseppe Bove	University of Rome 3
Eugenio Brentari	University of Brescia
Ida Camminatiello	University of Campania 'L. Vanvitelli'
Maurizio Carpita	University of Brescia
Carlo Cavicchia	Erasmus University Rotterdam
Roy Cerqueti	University of Rome 'La Sapienza'
Enrico Ciavolino	University of Salento
Corrado Crocetta	University of Bari 'Aldo Moro'
Antonello D'Ambra	University of Campania 'L. Vanvitelli'
Pierpaolo D'Urso	University of Rome 'La Sapienza'
Tonio Di Battista	University 'G. d'Annunzio' of Chieti-Pescara
Simone di Zio	University 'G. d'Annunzio' of Chieti-Pescara
Mauro Ferrante	University of Palermo
Francesca Fortuna	University of Rome 3
Michele Gallo	University of Naples 'Orientale'
Claudiu Herteliu	Bucharest University of Economic Studies
Salvatore Ingrassia	University of Catania
Ron Kenett	Samuel Neaman Institute & KPA Group
Michele La Rocca	University of Salerno
Rosaria Lombardo	University of Campania 'L. Vanvitelli'
Antonio Lucadamo	University of Sannio
Marica Manisera	University of Brescia
Andrea Marletta	University of Milan 'Bicocca'
Lucio Masserini	University of Pisa
Fabrizio Maturo	University Mercatorum
Matteo Mazziotta	ISTAT
Rodolfo Metulini	University of Bergamo
Stefania Mignani	University of Bologna
Enrico Ripamonti	University of Brescia

Silvia Salini	University of Milano
Pasquale Sarnacchiaro	University of Naples 'Federico II'
Violetta Simonacci	University of Naples 'Federico II'
Maurizio Vichi	University of Rome 'La Sapienza'
Mariangela Zenga	University of Milan 'Bicocca'

### **Organizing Committee**

<b>Chair:</b> Tonio Di Battista	University 'G. d'Annunzio' of Chieti-Pescara
Andrea Bucci	University 'G. d'Annunzio' of Chieti-Pescara
Alfredo Cartone	University 'G. d'Annunzio' of Chieti-Pescara
Marco Di Marzio	University 'G. d'Annunzio' of Chieti-Pescara
Marta Di Nicola	University 'G. d'Annunzio' of Chieti-Pescara
Simone Di Zio	University 'G. d'Annunzio' of Chieti-Pescara
Stefania Fensore	University 'G. d'Annunzio' of Chieti-Pescara
Lara Fontanella	University 'G. d'Annunzio' of Chieti-Pescara
Stefano Antonio Gattone	University 'G. d'Annunzio' of Chieti-Pescara
Luigi Ippoliti	University 'G. d'Annunzio' of Chieti-Pescara
Eugenia Nissi	University 'G. d'Annunzio' of Chieti-Pescara
Paolo Postiglione	University 'G. d'Annunzio' of Chieti-Pescara
Annalina Sarra	University 'G. d'Annunzio' of Chieti-Pescara
Pasquale Valentini	University 'G. d'Annunzio' of Chieti-Pescara
Adelia Evangelista (Secretariat)	University 'G. d'Annunzio' of Chieti-Pescara

## Editors

Andrea Bucci	University of Macerata
Alfredo Cartone	University 'G. d'Annunzio' of Chieti-Pescara
Adelia Evangelista	University 'G. d'Annunzio' of Chieti-Pescara
Andrea Marletta	University of Milan 'Bicocca'



Il Viandante - Copyright © 2023 Edizioni BACME S.r.l.s.  
Sede legale: Via Silvino Olivieri, 111 - 66100 Chieti (CH)

ISBN 979-12-803-3369-8

DOI 10.60984/978-88-94593-36-5-IES2023

<https://doi.org/10.60984/978-88-94593-36-5-IES2023>

All rights reserved.

This work is protected by copyright law.

All rights, in particular those relating to translation, citation, reproduction in any form, to the use of illustrations, tables and the software material accompanying the radio or television broadcast, the analogue or digital recording, to publication and dissemination through the internet are reserved, even in the case of partial use. The reproduction of this work, even if partial or in digital copy, is admitted only and exclusively within the limits of the law and is subject with the authorization of the publisher. Violation of the rules involves the penalties provided for by the law.

Sponsored by



# Preface

Statistical thinking, design and analysis play a crucial role in social life and are useful to society at large. Besides, promoting advanced methodological research is useful to facilitate the dissemination of ideas related to various fields of interest. For this purpose, experts in statistics, data analysis, data mining, statistical methods for decision making, machine learning and related methods come together to understand and analyse phenomena through data.

In line with this objective, the Statistics Group for the Evaluation and Quality of Services (SVQS; [www.svqs.it](http://www.svqs.it)) of the Italian Statistical Society (SIS) has been organizing the Innovation and Society (IeS) conference biennially since 2009, focusing on new developments and ideas in statistics applied to the evaluation and quality of public and private services, attracting national and international statisticians and data scientists. The meeting contributes to spot light on the main statistical approaches and methodologies for the evaluation of public services currently in use in different contexts, as well as to facilitate discussion on the impact of innovative statistical evaluation systems for these services, involving various economic and social policy actors.

The conference “Statistical Methods for Evaluation and Quality: Techniques, Technologies and Trends (T<sup>3</sup>)” recorded valuable contributions that are reported in this volume. The papers underscore how the growing availability of data has tasked social and economic actors, organizations, and researchers with the management and analysis of large volumes of unstructured and heterogeneous data. In recent years, many tools for both qualitative and quantitative models have been developed to better describe and understand complex systems and their underlying behaviors, and the papers reported in this volume bear witness to this.

Techniques, technologies and trends: the study of data complexity presents the potential to provide analyses with increased frequency and timeliness, accuracy and objectivity, and to define sustainable models. Traditional quantitative methods for capturing socioeconomic data have often shown limitations in their ability to examine underlying systems, and with the three ‘T’ just mentioned, the outlines of future developments are starting to emerge.

The volume reports 127 contributions in the following areas:

- Advanced statistical methods for pattern recognition
- Advances in statistical learning from high-dimensional data
- Data analysis for web sources
- Distance and depth-based statistical learning methods for robust data analysis

- Economics and environment
- Education and labour
- Inequalities in the labour market
- Innovations and challenges in official statistics
- Labour market: trends, perspectives and new challenges
- Methodological and applicative contributions for evaluating sustainable development
- Methodological developments and applications for the assessment of student competencies
- Networks data analysis: new perspectives and applications
- New advanced statistical methods for data science
- Recent advances in statistical learning and data analysis
- Statistical analysis and modeling of environmental pollution data
- Statistical methods and complexity for evaluation in finance
- Statistical methods and composite indicators for healthcare
- Statistical methods and models for land monitoring with spatio-temporal data
- Statistical methods for environmental monitoring and sustainability
- Statistical methods for the analysis of university student choices and academic performance
- Statistical methods for the assessment of transport services and sustainable emissions
- Statistical methods for education and educational services
- Statistics in sports
- Tourism and territory.

The Conference event attracted many contributions as well as numerous Authors, not just from Italy but also from abroad. Over the three-day meeting, the Community has the opportunity to witness some of the state-of-the arts, new trajectories, and methodological challenges in 24 solicited sessions, 7 sessions of free contributes, two round tables - organized by Maurizio Vichi and Matilde Bini respectively - and three keynotes sessions with Ron S. Kennet of Samuel Neaman Institute of Israel, Luigi D'Ambra of Federico II University of Naples, and the former Minister Enrico Giovannini from University of Tor Vergata.

*Organizers*

Paolo Mariani

*Chair of the Program Committee*

Tonio Di Battista

*Chair of Local Organizing Committee*

Maurizio Carpita

*Coordinator of the SVQS Group*



# Contents

<b>Plenary Session</b>	<b>1</b>
Befitting Cross Validation with Three Case Studies . . . . .	2
<i>Kenett R.S.</i>	
<b>Solicited Session SS1 - <i>Statistical analysis and modeling of environmental pollution data</i></b>	<b>3</b>
Assessing environmental quality by clustering a structural equation model based index: An application to European cities air pollution . . . . .	4
<i>Bottazzi Schenone M., Grimaccia E. and Vichi M.</i>	
Evaluating the nonlinear association between PM <sub>10</sub> and emergency department visits . . . . .	10
<i>Bucci A., Sanmarchi F., Santi L., Giostra F., Tubertini E., Rosa S., Nante N. and Golinelli D.</i>	
Estimating spatially varying Gaussian Graphical Models to unveil relationships among pollutants in the Red River Delta in Vietnam . . . . .	16
<i>Pronello N., Cucco A., Ignaccolo R. and Ippoliti L.</i>	
<b>Solicited Session SS2 - <i>Statistics in sports</i></b>	<b>22</b>
Clustering Athlete Performances in Track and Field Sports . . . . .	23
<i>Argiento R., Colombi A., Modotti L. and Montagna S.</i>	
A Cross-Country Analysis of Engagement in Physical Activity and Sport Practice Learnt from Eurobarometer Survey Data . . . . .	29
<i>Simone R.</i>	
Strong eras in male professional tennis . . . . .	35
<i>Breznik K., Candila V., Milekhina A. and Restaino M.</i>	
NonParametric Combination method for data analytics in basketball matches . . . . .	41
<i>Barzizza E., Biasetton N., Ceccato R., Disegna M. and Vezzosi G.</i>	
<b>Solicited Session SS3 - <i>Statistical methods for the analysis of university student choices and academic performance</i></b>	<b>47</b>
The influence of labor market conditions on students' career disruption: first insights from Italy . . . . .	48
<i>Usala C., Sulis I. and Porcu M.</i>	
Socio-economic aspects that may affect South-North students' mobility in Italy . . . . .	53
<i>Genova V.G. and Boscaïno G.</i>	
An analysis of student's performance in bachelor's degree . . . . .	59

<i>La Rocca M., Niglio M. and Restaino M.</i>	
An exploratory strategy for analyzing students' mobility data . . . . .	65
<i>Primerano I. and Giordano G.</i>	
<b>Solicited Session SS4 - <i>Statistics in football</i></b>	<b>71</b>
Community Detection in Sport Market Networks: The Case of Italian Professional Football . . . . .	72
<i>Randinelli R. and Ievoli R.</i>	
An Original Application to Football of PLS-SEM for the xG Model . . . . .	78
<i>Cefis M. and Carpita M.</i>	
Performance Assessment of Football Players and Combined Permutation Tests with application to Home-Field Advantage . . . . .	84
<i>Bonnini S.</i>	
A First Proposal of the Triad Census for Weighted Networks: an Application to Football . . . . .	90
<i>Randinelli R. and Palazzo L.</i>	
<b>Solicited Session SS5 - <i>Advances in statistical learning from high-dimensional data</i></b>	<b>96</b>
PCA approaches for vector functional time series . . . . .	97
<i>Aguilera A.M., Alonso F.J. and Acal C.</i>	
Conformal Prediction for Functional Kriging Models . . . . .	102
<i>Diana A., Romano E. and Adzic J.</i>	
Measuring Public-Private Connectedness in Financial Markets . . . . .	108
<i>Sánchez García J. and Cruz Rambaud S.</i>	
An original approach to anomalies in intertemporal choices through Functional Data Analysis: Theory and application for the study of Hikikomori syndrome . . . . .	112
<i>Martino R., Ventre V., Cruz Rambaud S. and Maturo F.</i>	
<b>Solicited Session SS6 - <i>Labour market: trends, perspectives and new challenges</i></b>	<b>118</b>
Enriching Job Vacancy Official Information with Online Job Advertisements: Chances and Limits . . . . .	119
<i>Lucarelli A. and Righi A.</i>	
Innovation in Management: towards the Open Manager . . . . .	126
<i>Bruttini P., Gallo M., Mariani P. and Menini T.</i>	
<b>Solicited Session SS7 - <i>Data analysis for web sources</i></b>	<b>132</b>
Enhancing SMEs default prediction with web-scraped data . . . . .	133
<i>Crosato L., Domenech J. and Liberati C.</i>	
Web data as enabler for informed decisions in Labour Market . . . . .	137
<i>Maggioni G.</i>	
The metaverse & luxury fashion brands: strategic communication exercise	141
<i>Forciniti A. and Zavarrone E.</i>	
Increasing the Geographical Granularity of Economic Indicators with Google Trends . . . . .	147

*Domenech J. and Marletta A.*

<b>Solicited Session SS8 - <i>Methodological and applicative contributions for evaluating sustainable development</i></b>	<b>153</b>
Evaluating sustainable development in EU countries through synthetic indicators . . . . .	154
<i>Alaimo L.S. and Cucci M.</i>	
Naples and tourism sustainability: A survey of citizens' perceptions . . . .	160
<i>Aria M., Pagliara F., D'Aniello L. and Della Corte V.</i>	
Modelling inequalities for sustainable development in Italy countries . . . .	166
<i>Musella M., Borrata G., Camminatiello I. and Lombardo R.</i>	
Food Security and Sustainability: A Science Mapping Analysis . . . . .	172
<i>Piscitelli A.</i>	
<b>Solicited Session SS9 - <i>Inequalities in the labour market</i></b>	<b>178</b>
Skill similarities across Italian regions: an analysis based on the online job advertisements . . . . .	179
<i>Kahlawi A., Buzzigoli L., Grassini L., Martelli C. and Giambona F.</i>	
Italian Labour Market reform and gender inequalities . . . . .	185
<i>Marini C. and Nicolardi V.</i>	
Intergenerational transmission of disadvantages in the Italian labour market: evidence from AD-SILC data . . . . .	191
<i>Busetta A., Fabrizi E., Ragozini G. and Sulis I.</i>	
<b>Solicited Session SS10 - <i>Statistical methods and complexity for evaluation in finance</i></b>	<b>197</b>
Financial networks resilience and shocks propagation . . . . .	198
<i>Cerqueti R., Cinelli M., Ferraro G. and Iovanella A.</i>	
How the choice of one parameter impacts the numerical stability of the efficient frontier . . . . .	204
<i>Fassino C. and Uberti P.</i>	
Dynamic shrinkage for minimum variance combination of forecasts . . . . .	210
<i>Mattera R.</i>	
Exploring the perception of the gender issue of Italian female entrepreneurs	217
<i>Castellano R., Riccioni J. and Rinaldi A.</i>	
<b>Solicited Session SS11 - <i>Networks data analysis: new perspectives and applications</i></b>	<b>223</b>
Describing Italian mobility trajectories in higher education . . . . .	224
<i>Genova V.G., Giordano G., Ragozini G. and Vitale M.P.</i>	
Collaboration networks: methodological issues and updated empirical evidence on Italian statisticians . . . . .	230
<i>De Stefano D., Fabbrucci Barbagli A.G., Santelli F. and Zaccarin S.</i>	
Mapping Ashtma Complexity with Graph Theory: an Integrative Approach	236
<i>Cucco A., Simpson A., Murray C., Fontanella S. and Custovic A.</i>	
Investigating the patterns of Italian internal mobility: a network analysis at provincial level . . . . .	242

*Sarra A., D'Ingiullo D., Evangelista A., Nissi E., Quaglione D. and Di Battista T.*

<b>Solicited Session SS12 - <i>Innovations and challenges in official statistics</i></b>	<b>248</b>
Formal and informal networks of care for the elderly: regional profiles compared . . . . .	249
<i>Sicuro L., Tucci D. and Coniglio R.</i>	
Gender Gap: a multidimensional approach . . . . .	255
<i>Acampora C., Fusco D., Liguori M.A. and Pagliuca M.M.</i>	
Using Whatsapp in Official Statistics: a New tool for managing the Agriculture Census . . . . .	259
<i>Fabi C.</i>	
<b>Solicited Session SS13 - <i>Statistical methods and composite indicators for healthcare</i></b>	<b>265</b>
Longitudinal composite indicators to measure the quality of health services	266
<i>Crocetta C., Antonucci L., Cataldo R. and Mazza R.</i>	
Past and Future of Doctor-Patient Communication . . . . .	272
<i>Tedesco N., Zavarrone E. and Forciniti A.</i>	
Network Analysis approach to customer satisfaction and service quality detection: an application to health-care services . . . . .	277
<i>Crocetta C., Grassia M.G., Marino M., Mazza R., Simonacci V. and Stavolo A.</i>	
A project evaluation study on multiset Likert scale data . . . . .	283
<i>Simonacci V., Marino M., Grassia M.G. and Gallo M.</i>	
<b>Solicited Session SS14 - <i>Distance and depth-based statistical learning methods for robust data analysis</i></b>	<b>289</b>
Robust distance-based predictive models . . . . .	290
<i>Boj E., Grané A. and Parron D.</i>	
Data depth for mixed-type data through multidimensional scaling. An application to biological age imputation . . . . .	294
<i>Cascos I., Grané A. and Qian J.</i>	
A compared protocol to improve clustering procedures . . . . .	298
<i>Grané A., Riani M. and Salini S.</i>	
Robust diagnostics for Linear Mixed Models with the Forward Search . . .	304
<i>Corbellini A., Grossi L. and Laurini F.</i>	
<b>Solicited Session SS15 - <i>Advanced statistical methods for pattern recognition</i></b>	<b>310</b>
Unsupervised classification of NPLs recovery curves . . . . .	311
<i>Carleo A. and Rocci R.</i>	
Living alone in Italian municipalities . . . . .	317
<i>Vellucci P., Benassi F., Naccarato A. and Gallo G.</i>	
Supervised learning from high-dimensional data through dynamic updating of functional classification rules . . . . .	323

<i>Maturo F., Fortuna F. and Di Battista T.</i>	
Assessing the effectiveness of coordination among public authorities in cohesion expenditure . . . . .	329
<i>Coco G., Monturano G. and Resce G.</i>	
<b>Solicited Session SS16 - <i>Recent advances in statistical learning and data analysis</i></b>	<b>335</b>
A Predictive Functional Principal Component Analysis of Well-Being Data	336
<i>Marcis L., Pagliarella M.C. and Salvatore R.</i>	
Detecting the partition in the extended hierarchy of a dendrogram: an application on biomedical data . . . . .	343
<i>PolICASTRO V., Palazzo L. and VISTOCCO D.</i>	
Concordance measure for rankings . . . . .	350
<i>Bissiri P.G. and Nai Ruscone, M.</i>	
Quadratic discriminant scoring for selecting clustering solutions . . . . .	355
<i>Coraggio L. and Coretto P.</i>	
<b>Solicited Session SS17 - <i>Statistical methods for education and educational services</i></b>	<b>361</b>
Association between INVALSI scores and students' mobility in Italy: a preliminary analysis . . . . .	362
<i>Bacci S., Bertaccini B., Lombardi G. and Tocchioni V.</i>	
Modelling Responses and Response Times: an application to Mathematics INVALSI data . . . . .	368
<i>Bungaro L., Desiderati R. and Mignani S.</i>	
Latent potential outcomes: An analysis of the effects of programs aimed at improving students' non-cognitive skills . . . . .	374
<i>Pennoni F., Bartolucci F. and Vittadini G.</i>	
Cognitive Skills and Non Cognitive Skills to Analyze School and Students Performances . . . . .	380
<i>Vittadini G.</i>	
<b>Solicited Session SS18 - <i>Statistical methods for the assessment of transport services and sustainable emissions</i></b>	<b>386</b>
Sustainability assessment of urban transport by an LCA comparison on different technologies vehicles . . . . .	387
<i>Della Ragione L. and Meccariello G.</i>	
Passenger comfort prediction via time-series classification . . . . .	393
<i>Vanacore A., Pellegrino M.S. and Ciardiello A.</i>	
A statistical model to analyse driving behavior: a case study . . . . .	399
<i>Rodia G., Sarnacchiaro P. and Acciarino V.</i>	
Aggregating judgments in non negotiable group decisions in transport system	403
<i>Amenta P. and Lucadamo A.</i>	
<b>Solicited Session SS19 - <i>New advanced statistical methods for data science</i></b>	<b>408</b>

A unified framework for two-dimensional clustering on preference-approvals: an analysis of Eurobarometer data . . . . .	409
<i>Albano A., Sciandra M. and Plaia A.</i>	
Pandemic Data Quality Modelling: A Bayesian Approach . . . . .	415
<i>Ferrari L., Manzi G., Micheletti A., Nicolussi F. and Salini S.</i>	
Explainable AI for Peer-to-Peer Credit Risk Management . . . . .	421
<i>Babaei G., Pagnottoni P. and Do T. T.</i>	
Tackling misclassification in surveys about undeclared work via the EM algorithm . . . . .	427
<i>Arezzo M.F., Guagnano G. and Vitale D.</i>	
<b>Solicited Session SS20 - <i>Tourism, territory and data analysis</i></b>	<b>432</b>
Tourism, sustainability, and territorial impact: an input-output analysis . .	433
<i>Garau G. and El Meligi A.K.</i>	
The Impact of Big Data in Tourism . . . . .	439
<i>Ciuffreda R., Choedon C. and Simonetti B.</i>	
Neural network-based prediction of domestic tourists' length of stay in Italy	443
<i>Antolini F. and Cesarini S.</i>	
The management of cultural heritage in contexts of undertourism: a model for assessing the economic sustainability of public-private partnerships	450
<i>Calabrò F.</i>	
<b>Solicited Session SS21 - <i>Statistical methods and models for land monitoring with spatio-temporal data</i></b>	<b>457</b>
Geo-referenced data and complex networks for measuring road accident risk	458
<i>Cantaluppi G., Clemente C., Della Corte F. and Zappa D.</i>	
A comparison of geospatial models for car crash risk . . . . .	464
<i>Cantaluppi G., Giardino G. and Zappa D.</i>	
Geostatistical modelling of livestock-related PM <sub>2.5</sub> pollution and scenario analysis for policymakers - Work in progress . . . . .	470
<i>Fassò A., Rodeschini J., Fusta Moro A. and Finazzi F.</i>	
Functional clustering methods for space-time big data from mobile phone networks . . . . .	476
<i>Perazzini S., Metulini R. and Carpita M.</i>	
<b>Solicited Session SS22 - <i>Methodological developments and applica- tions for the assessment of student competencies</i></b>	<b>483</b>
Modeling the main drivers of mathematical literacy of school-leaving stu- dents. Some evidence from the Invalsi tests . . . . .	484
<i>Davino C., Palumbo F., Romano R. and Vistocco D.</i>	
Educational Data Mining: clustering students' performance over time . . .	490
<i>Taraborrelli G. and Farnè M.</i>	
The nexus of cultural capital with participation in early childhood education	496
<i>Ripamonti E.</i>	
High- and Low-Performing students and future career: a gender and social issue . . . . .	502
<i>Falzetti P. and Ricci R.</i>	

<b>Solicited Session SS23 - <i>Statistical methods for environmental monitoring and sustainability</i></b>	<b>508</b>
Clustering spatial data through optimal transports . . . . .	509
<i>Balzanella A. and Verde R.</i>	
New interpretative insights for environmental air quality by means of FDA	514
<i>Terzi S., Naccarato A. and Fortuna F.</i>	
A Bayesian State-Space Model to Mitigate Unmeasured Confounding . . .	520
<i>Zaccardi C., Valentini P. and Ippoliti L.</i>	
Mining social media data for damage assessment in environmental disasters	526
<i>del Gobbo E., Cafarelli B., Ippoliti L. and Fontanella L.</i>	
<b>Solicited Session SS24 - <i>Satisfaction and behavior in tourism</i></b>	<b>531</b>
The evaluation of the hotel stay through a new development of correspondence analysis coping with ordinal variables . . . . .	532
<i>D'Ambra A. and Amenta P.</i>	
Assessing the role of knowledge and authenticity in the formation of attendee loyalty at cultural festivals . . . . .	536
<i>Rivetti F., Lucadamo A. and Rossi C.</i>	
Residents' Opinions and Perceptions of Tourism Development in the Historic City of Matera . . . . .	542
<i>Sarnacchiaro P., Di Gennaro R. and Di Taranto E.</i>	
Exploring tourism at religious sites: The case of Assisi . . . . .	547
<i>Rivetti F., Dini M. and Splendiani S.</i>	
<b>Session of free contributes SFC1 - <i>Education and labour</i></b>	<b>552</b>
Local concordance among the items of questionnaires on student's opinion (OPIS) . . . . .	553
<i>Terzi S. and Petrarca F.</i>	
High School Proficiency of Future University Students: An Analysis based on INVALSI Data . . . . .	560
<i>Santelli F., Di Credico G. and Di Caterina C.</i>	
Employment vulnerability of immigrants in the labour market – Does origin matter? . . . . .	566
<i>Bittaye M.</i>	
The effect of pricing policies on students' use of university canteens . . . .	572
<i>Masserini L., Bini M. and Lorenzoni V.</i>	
Gig workers' identikit . . . . .	576
<i>Zavarrone E. and Forciniti A.</i>	
<b>Session of free contributes SFC2 - <i>Tourism and territory</i></b>	<b>582</b>
Multi-source approach for statistics in tourism sector . . . . .	583
<i>Bianchino A., Fusco D., Giordano P., Liguori M.A. and Summa D.</i>	
Statistical analysis of tourism sustainability in Campania: post Covid-19 review . . . . .	589
<i>Giacalone M., Basile V. and Bellucci M.</i>	
Investigating recent changes in dietary behavior . . . . .	596
<i>D'Uggento A., d'Ovidio F.D., Toma E. and Onorati M.G.</i>	



Depopulation in the Abruzzo municipalities . . . . .	603
<i>Carulli A.L., Di Spalatro D. and Valentini A.</i>	
The Productions System of Inland Areas . . . . .	607
<i>Madia Carucci A.M. and Regano A.</i>	
<b>Session of free contributes SFC3 - <i>Social issues</i></b>	<b>613</b>
Modelling the gender gap in youth mortality with an Age-Period-Cohort analysis . . . . .	614
<i>Lanfiuti Baldi G. and Nigri A.</i>	
Random forest for classifying odor emission sources . . . . .	620
<i>Distefano V., Palma M., De Iaco S. and Mazuruse G.</i>	
An Experimental Annotation Task Investigating Annotator Agreement Within a Misogynistic Dictionary and Corpus . . . . .	626
<i>Tontodimamma A., Ignazzi E., Anzani S., Fontanella L. and Di Zio S.</i>	
Statistical analysis of COVID19 impact on Italian mortality . . . . .	632
<i>Franchetti G. and Politano M.</i>	
Measuring multidimensional deprivation using objective and subjective data: an application of the Voronoi ranking method . . . . .	638
<i>Ciommi M., Mariani F., Polinesi G. and Recchioni M.C.</i>	
<b>Session of free contributes SFC4 - <i>Trends</i></b>	<b>644</b>
The role of big data analytics in circular supply chains: A bibliometric analysis . . . . .	645
<i>Khan F. and Rapposelli A.</i>	
Estimation of the ranking of incentive policies for the adoption of 4.0 tech- nologies . . . . .	652
<i>Bonnini S. and Borghesi M.</i>	
Risk Management and Future Scenarios. A proposal based on a mixed- method approach . . . . .	658
<i>Di Zio S., Bolzan M., Marozzi M. and Scioni M.</i>	
Explainable artificial intelligence (XAI) through artificial intelligence from a human in the loop (HITL) perspective: an interview with ChatGPT	664
<i>Santarcangelo V., Lamacchia A., Vitullo S., Di Lecce M. and Giacalone M.</i>	
Relevance in official statistics: information needs, satisfaction with data quality, some results and future perspectives . . . . .	670
<i>Loporcaro M.F.</i>	
<b>Session of free contributes SFC5 - <i>Economic issues</i></b>	<b>676</b>
Evolutionary trends of start-ups in Italy: a case study . . . . .	677
<i>Duttilo P., Caruso G., Iannone B. and Gattone S.A.</i>	
Permanent establishments and efficiency analysis with global enterprises .	682
<i>Frenda A. and Sepe E.</i>	
Techniques and constructs in some recent market and organizational research	685
<i>Sciascia I.A.</i>	
The value of buildings in the Italian general government balance sheet . . .	689
<i>Santoro P. and Regano A.</i>	



Linear and nonlinear factors affecting default risk in the peer-to-peer lending market . . . . .	695
<i>Giordano F., Milito S. and Parrella M.L.</i>	
<b>Session of free contributes SFC6 - <i>Methodological issues</i></b>	<b>701</b>
Macroeconomic Time Series Classification by Nonparametric Trend Estimation . . . . .	702
<i>Feo G., Giordano F., Niglio M. and Parrella M.L.</i>	
A Normalization Method for Space-time Analysis of Evaluation and Quality Indicators . . . . .	709
<i>Mazziotta M. and Pareto A.</i>	
Unveiling Latent Structures: exploring Multidimensional IRT Models using Dirichlet Process Mixtures . . . . .	713
<i>Valentini P., Fontanella S. and Fontanella L.</i>	
On a technique to detect accounting data manipulation . . . . .	718
<i>Passamonti C.</i>	
<b>Session of free contributes SFC7 - <i>Economics and environment</i></b>	<b>724</b>
Determinants of Water Conservation Behaviour and Spatial Heterogeneity in their Coefficients . . . . .	725
<i>Mammadli R. and Gigliarano C.</i>	
Modeling the economic burden of grass pollen allergoid immunotherapy . .	731
<i>Bilancia M. and Di Bona D.</i>	
Swine fever in Liguria: who does pay for economic losses? A causal analysis	735
<i>Baggetta C., Cavalletti B. and Corsi M.</i>	
Measuring Multidimensional Poverty of the Italian Regions in the era of COVID-19 . . . . .	741
<i>Chelli F.M.C., Ciommi M., Gigliarano C. and Polinesi G.</i>	
Drivers of inflation: relationships changing over time . . . . .	747
<i>Sokolenko O., Palumbo A., Fortuna F., Naccarato A. and Marie J.</i>	
Environmental accounting and sustainable cities: an explorative bibliometric-based literature analysis . . . . .	753
<i>Kaur A., Javed A., D'Andreamatteo A. and Rapposelli A.</i>	

# Investigating the patterns of Italian internal mobility: a network analysis at provincial level

## *La mobilità interna italiana: un'analisi di rete a livello provinciale*

Annalina Sarra, Dario D'Ingiullo, Adelia Evangelista, Eugenia Nissi, Davide Quaglione and Tonio Di Battista

**Abstract** This study employs network analysis to investigate internal migration patterns in Italy from 2002 to 2018. Using filtering methods, the observed dense network was reduced to its essential components. The extracted backbone's characteristics were then examined, including measures of structural balance and community detection. Our findings suggest that the majority of signed links in the network exhibit reciprocity, and that positive relationships between geographically close-by cities and their shared locations at greater distances contribute to a structural balance. Additionally, a community analysis revealed that geography play a crucial role in defining communities within the network.

**Abstract** *Questa ricerca utilizza l'analisi delle reti per indagare i modelli di migrazione interna in Italia dal 2002 al 2018. Utilizzando metodi di filtraggio, la fitta rete osservata è stata ridotta alle sue componenti essenziali. Le caratteristiche del backbone estratto sono state quindi esaminate, inclusi i parametri di equilibrio strutturale e la rilevazione delle comunità. I nostri risultati suggeriscono che la maggior parte dei collegamenti con segno nella rete mostrano reciprocità, e che le relazioni positive tra le città geograficamente vicine e le loro destinazioni comuni a distanze maggiori contribuiscono ad un equilibrio strutturale. Inoltre,*

---

Annalina Sarra  
University "G.d' Annunzio" of Chieti-Pescara, e-mail: annalina.sarra@unich.it

Dario D'Ingiullo  
University "G.d' Annunzio" of Chieti-Pescara, e-mail: dario.dingiullo@unich.it

Adelia Evangelista  
University "G.d' Annunzio" of Chieti-Pescara, e-mail: adelia.evangelista@unich.it

Eugenia Nissi  
University "G.d' Annunzio" of Chieti-Pescara, e-mail: eugenia.nissi@unich.it

Davide Quaglione  
University "G.d' Annunzio" e-mail: davide.quaglione@unich.it

Tonio Di Battista  
University "G.d' Annunzio" e-mail: tonio.dibattista@unich.it

Sarra A., D’Ingiullo D., Evangelista A., Nissi E., Quaglione D. and Di Battista T.

*un’analisi delle comunità ha indicato che la geografia svolge un ruolo cruciale nella definizione delle comunità all’interno della rete.*

**Key words:** Internal migration, Italy, Network analysis, Filtering methods, Community detection.

## 1 Introduction

Internal migration – that is the share of individuals who permanently transfer their place of residence within the national boundaries – affecting the spatial redistribution of people and their characteristics, is recognized to be an important mechanism able to influence several socio-economic outcomes [4]. Hence, investigating the structure and dynamics of the internal migrations should occupy a central stage on academic and policy debate in order to develop and put in place sound policies for a given economy. Studying internal migration patterns is crucial in countries like Italy, our area of focus, where significant territorial imbalances exist. It allows for a better understanding of the dynamics of population movements within the country’s economy, highlighting areas of growth and areas that may require additional support. The literature strand on the determinants of Italian migration has demonstrated, in fact, how better socio-economic and institutional characteristics of the North induce a population redistribution towards this area, which in turn can further improve its already better structural characteristics (if these flows are endowed with a high level of education) [5]. By considering the consequences, in fact, the growing age and skill selectivity that characterizes the internal flows, contributing to spread skills and knowledge across regions, makes internal migration a crucial growth enhancing mechanism in the host economy [3, 1]. Therefore, this selectivity in migration contributes to feed a vicious circle to the detriment of the already backward regions of the *Mezzogiorno*. In this frame, an important contribution stems from an analytical tool largely adopted in several economic and non-economic fields: the network analysis. This methodology is particularly useful to provide an immediate representation of the relationships and functional linkages that exist among regions of a given economy, by combining measures of network connectivity with measures of network shape [2]. Building upon this, we extend the commonly used network analysis methods by incorporating a filtering technique that extracts both positive and negative migration links (i.e., a signed network) and by utilizing machine learning methods to analyze the low-dimensional representation of the migration system. This approach, as proposed by [7], allows for a more comprehensive analysis of the migration patterns and dynamics. In doing this, the present article makes use of a unique database that consists of bilateral migration flows among Italian regions (NUTS-3) collected by the Italian National Institute of Statistics (ISTAT) in the Migratory movements of resident population – registrations and cancellations to the registry office. The reminder of the article is structured as follows: Section 2 illus-

trates the data collection and the methodology. In Section 3 the results are presented and discussed.

## 2 Data and methods

### 2.1 Data

The bilateral migration flows among the 103 Italian provinces for the years 2002-2018 are collected by the ISTAT in the Migratory movements of resident population, registrations and cancellations to the registry office. For each year, by considering all the possible migration flows among the 103 Italian provinces, the resulting bilateral matrix of interprovincial movements is a square matrix that has a dimension of 103 x 103. Obviously, by excluding the main diagonal, which contains only zeros by definition (i.e. we do not consider the intra-provincial mobility), the total number of observations for a single year is equal to 10,506.

### 2.2 Methodology: Network analysis

Based on the approach outlined by [6], we implemented a four-step methodology to analyze the structure and dynamics of internal migration within Italy at the provincial level. Our first step involved constructing a directed and weighted network to model migration in-flows and out-flows between cities. Specifically, we created a distinct network for each year, in which links represented the movement of individuals from an origin city to a destination city, with the weight of each link indicating the number of people migrating. Formally, we represented the directed and weighted network as  $G(V, E, W)$ , where  $|V|$  is the set of nodes denoted by  $i, j, \dots$ ,  $|E|$  is the set of directed edges, and  $W$  is the weighted adjacency matrix. The network consisted of  $n$  nodes and  $m$  edges, with  $W_{ij}$  representing the number of migrants moving from node  $i$  to node  $j$ , and  $W_{ii}$  being equal to 0. In the second step, we aimed to extract relevant and significant information from the initial networks by simplifying and reducing them into a sparser format that preserves sufficient structural information for efficient analysis. This resulted in the creation of a meaningful signed network, denoted by  $\hat{G} := (\hat{V}, \hat{A}, \hat{W})$ , where the sparse adjacency matrix  $\hat{A}$  typically only contains values of -1, 0, or 1. The sparse adjacency matrix is used to represent the links between nodes in the network, with a value of -1 indicating a negative link (an inhibitory relationship), a value of 1 indicating a positive link (an excitatory relationship), and a value of 0 indicating the absence of a link. The corresponding weight matrix,  $\hat{W}$ , is optional and only assigned to non-zero values in  $\hat{A}$ . If  $\hat{A}_{ij} = 0$ , then  $\hat{W}_{i,j}$  is also 0, otherwise,  $\hat{W}_{i,j}$  has the same sign as  $\hat{A}_{i,j}$ . This process involves information filtering, also known as “backbone extraction” or “network sparsifica-

tion,” which entails removing unnecessary or weak links in a network to extract its most important and informative connections. Various methods, including filtering techniques, can be used to achieve this “sparsification”. We utilized a recent dense network filtering method proposed by [6] in 2021 to extract both positive and negative links. This approach relies on a null model that considers the in-strength and out-strength of nodes to estimate expected link weights and identify links with significantly different weights as positive or negative. The method can extract the network’s signed backbone at a desired level of sparsity, based on the statistical significance of the links (via its significance filter) or the intensity of the links (via its vigor filter). The vigor filter is a lift-based measure that ranges from -1 to 1 and can be used to determine the strength of the links. By using this method, researchers can analyze important connections in a network and gain insights into its structure and function. Our subsequent objective was to create a concise representation of nodes by eliminating irrelevant information. This involved estimating an embedding matrix  $\mathbf{Z}$  with dimensions  $(n \times d)$ , where  $n$  is the number of nodes,  $d$  is the embedding size, and  $z_i$  represents the embedding vector of node  $i$ . Our goal was to find a relation between  $f(i, j)$  and  $g(z_i, z_j)$ , where  $f$  is a similarity function in the observed network and  $g$  is a similarity function in the latent space. We chose cosine similarity as the similarity function, given that the input data’s similarity range (i.e., edge weights in the signed backbone) is between -1 and 1. To complement the analysis of the migration system using lower dimensions, we implemented established clustering methods to identify community structures in the migration networks.

### 3 Results and discussion

In this section, we present the results of our study on migration patterns. In terms of the total number of migrants, we record a slow but often steady decrease in the number of people who have moved. The examination of the migration flows within the original network indicates that exists a reduced number of links that possess significantly greater weights. Additionally, the distribution of node strengths is characterized by a high degree of heterogeneity, with a few cities accounting for a disproportionate volume of migration. It is worth noting that, in the case of Italy’s internal migration, there are migration flows between all pairs of cities in both directions, resulting in a network that is inherently dense. This highlights the need to filter out insignificant links. To achieve this, we employed information filtering methods that took into account the network’s multiscale nature, thereby enabling the extraction of a sparse backbone. After analyzing the backbone’s characteristics, it becomes apparent that most signed links in the network demonstrate reciprocity, as they are reciprocated by links bearing the same sign. Even when the size of the backbone is considerably large, conflicting links are rare. We also evaluated whether the extracted network exhibits a structural balance or weak structural balance (Fig.1). Our results indicate the presence of a structural balance among the nodes in our networks, which may be attributed to positive relationships between geographically

close-by cities and their shared destinations at greater distances. Fig. 2 displays the spatial network of Italy’s internal migration backbone in 2018.

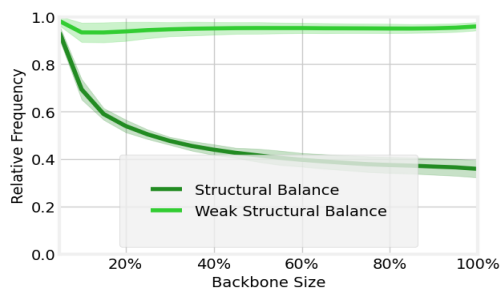


Fig. 1 Characteristics of migration backbones: structural balance

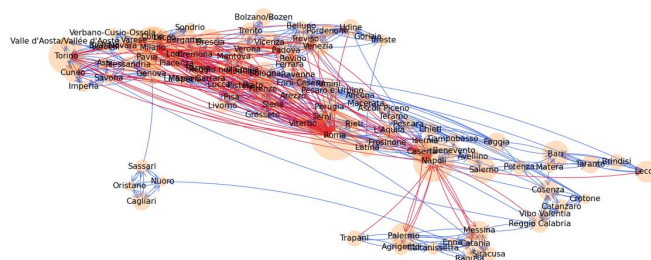
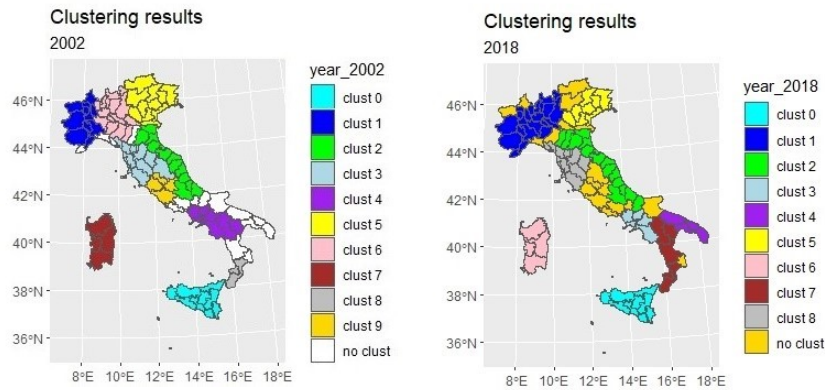


Fig. 2 Spatial network of migration backbone in 2018

Most of the positive links that extend beyond the local level in the network are oriented in a South-North direction. One possible explanation for this trend is that long-distance migrations typically occur between areas with significant economic hubs and activity centers that in Italy are located in the northern regions. We conclude our investigation with a community analysis. Figure 3 displays the results of the density-based clustering for both 2002 and 2018, using different colors for

each cluster. Our analysis indicates that geographical proximity plays a major role in defining communities. The results suggest that areas that are closer to each other tend to be clustered together, regardless of other factors such as population density or economic activity. This finding has important implications for understanding patterns of regional development and social interaction.



**Fig. 3** Density-based clusters

## References

1. Basile, R., Girardi, A., Mantuano, M., Russo, G.: Interregional Migration of Human Capital and Unemployment Dynamics: Evidence from Italian Provinces. *Ger. Econ. Rev.*, 1-30 (2018) doi:10.1111/geer.12172
2. Carvalho, R. C. de, Charles-Edwards, E.: The evolution of spatial networks of migration in Brazil between 1980 and 2010. *Popul. Space Place*, 26(7) e2332 (2020) doi:10.1002/psp.2332
3. Dotti, N. F., Fratesi, U., Lenzi, C., Percoco, M.: Local labour market conditions and the spatial mobility of science and technology university students: evidence from Italy. *Rev.Reg. Res.* 34(2), 119-137 (2014) doi: 10.1007/s10037-014-0088-y
4. Etzo, I.: Internal migration: a review of the literature. MPRA Paper Number. 8783 (2008)
5. Etzo, I.: The determinants of the recent interregional migration flows in Italy: A panel data analysis. *J. Reg. Sci.* 51(5), 948-966 (2011) doi:10.1111/j.1467-9787.2011.00730.x
6. Gürsoy, F., Badur, B.: Extracting the signed backbone of intrinsically dense weighted networks. *J.Complex Netw.* 9(5):cnab019.(2021) doi: 10.1093/comnet/cnab019
7. Gürsoy, F., Badur, B.: Investigating internal migration with network analysis and latent space representations: an application to Turkey. *Soc. Netw. Anal. Min.* 12(1), 1-16 (2022) doi:10.1007/s13278-022-00974-w