



IES2025 Bressanone

Innovation & Society: Statistics and Data Science for Evaluation and Quality

SVQS

Research Group of Statistics for the Evaluation and Quality in Services
of the Italian Statistical Society (SIS)

Innovation & Society

Statistics and Data Science for Evaluation and Quality

Editors: Giovanna Boccuzzo, Enrico Bovo,
Marica Manisera, Luigi Salmaso

University of Padova
Casa della Gioventù
Brixen – Bressanone
June 25-27, 2025

cleup

First edition: June 2025

ISBN 978 88 5495 849 4

© IES 2025 - Statistical Methods for Evaluation and Quality

2025 Cleup sc

“Coop. Libreria Editrice Università di Padova”

via G. Belzoni 118/3, Padova (t. +39 049 8753496), Italy

www.cleup.it

www.facebook.com/cleup

All rights reserved.

No part of this book may be translated or copied in whole or in part without written permission of the authors, except for brief excerpts in connection with reviews or scholarly analysis.

**IES 2025 - Innovation & Society: Statistics
and Data Science for Evaluation and Quality**

BOOK OF SHORT PAPERS

Editors: Giovanna Boccuzzo, Enrico Bovo,
Marica Manisera, Luigi Salmaso

Book of Short Papers

2025 Conference of the 12th Scientific Meeting of the Statistics for
the Evaluation and Quality of Services Group of the Italian
Statistical Society (SVQS)

IES 2025 - Innovation & Society: Statistics and Data Science for
Evaluation and Quality University of Padova

Casa della Gioventù, Brixen - Bressanone
June 25-27, 2025

Program Committee

Chair: Marica Manisera	University of Brescia
Ana María Aguilera	University of Granada - Sociedad de Estadística
Rosa Arboretti Giancristofaro	University of Padova
Francesca Bassi	University of Padova
Matilde Bini	European University of Roma (ISBIS)
Eugenio Brentari	University of Brescia
Mario Bolzan	University of Padova
Ida Camminatiello	University of Campania 'L. Vanvitelli'
Maurizio Carpita	University of Brescia
Carlo Cavicchia	Erasmus University Rotterdam
Giulia Cavrini	Free University of Bozen-Bolzano
Roy Cerqueti	University of Rome 'La Sapienza'
Enrico Ciavolino	University of Salento
Elisa Cisotto	Free University of Bozen-Bolzano
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 of Chieti-Pescara 'G. D'Annunzio'
Simone Di Zio	University of Chieti-Pescara 'G. D'Annunzio'
Marta Disegna	University of Padova
Michele Gallo	University of Naples 'L'Orientale'
Claudiu Herteliu	Bucharest University of Economic Studies
Salvatore Ingrassia	University of Catania
Michele La Rocca	University of Salerno
Christophe Ley	University of Luxembourg (President of the Luxembourg Statistical Society)
Caterina Liberati	University of Milano 'Bicocca' (ISBIS)
Rosaria Lombardo	University of Campania 'L. Vanvitelli'
Antonio Lucadamo	University of Sannio
Ambra Macis	University of Brescia
Paolo Mariani	University of Milano 'Bicocca'
Antonella Massari	University of Bari
Fabrizio Maturo	University Mercatorum
Matteo Mazziotta	ISTAT
Fatmir Memaj	University of Tirana
Tullio Menini	University of Naples 'L'Orientale'
Rodolfo Metulini	University of Bergamo
Stefania Mignani	University of Bologna

Fabrizio Ruggeri	CNR - IMATI
Silvia Salini	University of Milano
Pasquale Sarnacchiaro	University of Naples 'Federico II'
Kozeta Sevrani	University of Tirana and INSTAT
Violetta Simonacci	University of Naples 'Federico II'
Rosaria Simone	University of Naples 'Federico II'
Anna Simonetto	University of Brescia
Amalia Vanacore	University of Naples 'Federico II'
Vedat Verter	Queen's University Smith School of Business
Marika Vezzoli	University of Brescia
Maurizio Vichi	University of Rome 'La Sapienza'
Emma Zavarrone	IULM University

Local Organizing Committee

Co-Chair: Giovanna Boccuzzo	University of Padova
Co-Chair: Luigi Salmaso	University of Padova
Erika Banzato	University of Padova
Elena Barzizza	University of Padova
Pietro Belloni	University of Padova
Nicolò Biasetton	University of Padova
Enrico Bovo	University of Padova
Riccardo Ceccato	University of Padova
Pierdomenico Duttilo	University of Padova
Alberto Molena	University of Padova
Maurizio Nicolaio	University of Padova
Manuela Scioni	University of Padova
Margherita Silan	University of Padova
Francesco Trolese	University of Padova

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'

SVQS Thematic Area Leads

Matilde Bini	European University of Rome
Eugenio Brentari	University of Brescia
Maurizio Carpita	University of Brescia
Carlo Cavicchia	Erasmus University Rotterdam
Roy Cerqueti	University of Rome 'La Sapienza'
Enrico Ciavolino	University of Salento
Antonello D'Ambra	University of Campania 'L. Vanvitelli'
Tonio Di Battista	University of Chieti-Pescara 'G. D'Annunzio'
Michele Gallo	University of Naples 'L'Orientale'
Salvatore Ingrassia	University of Catania
Marica Manisera	University of Brescia
Paolo Mariani	University of Milan 'Bicocca'
Stefania Mignani	University of Bologna
Pasquale Sarnacchiaro	University of Naples 'Federico II'

Editors

Giovanna Boccuzzo

University of Padova

Enrico Bovo

University of Padova

Marica Manisera

University of Brescia

Luigi Salmaso

University of Padova

Preface

Statistical thinking and data-driven methodologies are essential to understanding contemporary challenges and driving innovation in several sectors. Since 2009, the Innovation and Society (IES) biennial conference, organized by the Group of Statistics for the Evaluation and Quality of Services (SVQS) of the Italian Statistical Society (SIS), has served as a key venue for advancing research in statistical methods applied to evaluation and quality.

The 12th edition, titled “Innovation & Society: Statistics and Data Science for Evaluation and Quality” took place from June 25 to 27, 2025, at the Bressanone-Brixen campus of the University of Padova. This year’s conference reinforced the role of statistics and data science as key tools for evaluating complex phenomena and ensuring quality in a variety of fields, including education, healthcare, environment, public administration, finance, tourism, sports, and beyond.

The scientific program was particularly rich, featuring 3 plenary lectures, 44 invited sessions, 6 contributed sessions, one special plenary lecture and a round table co-organized by The Italian (ISTAT) and Albanian (INSTAT) National Statistical Institutes. In total, 199 contributions were presented within invited and contributed sessions. The event also saw the highest international participation in the history of IES, with contributions from the Austrian and Luxembourg Statistical Societies, among others.

Three distinguished keynote speakers enriched the program with plenary lectures: Fabrizio Ruggeri, Christophe Ley, Werner Müller. In addition, a special plenary session featured two prominent invited talks by Fortunato Pesarin and Narayanaswamy Balakrishnan, further enhancing the depth and breadth of the scientific discussions.

The conference also hosted the Young Statisticians Award, which attracted 36 high-quality submissions, showcasing the energy and creativity of early-career researchers working at the intersection of statistics and real-world evaluation problems.

The short papers in this volume highlight the wide spectrum of topics discussed at IES 2025, and testify to the evolving role of statistics and data science in producing knowledge and supporting decision-making in complex systems.

We are sincerely grateful to all authors, session chairs and discussants, session organizers, reviewers, and participants. We also warmly thank our institutional partners for their support: Italian Statistical Society (SIS), SVQS Group, DMS StatLab of the University of Brescia, Departments of Statistical Sciences and of Managements and Engineering of the University of Padova, Österreichische Statistische Gesellschaft (Austrian Statistical Society), Interna-

tional Society for Business and Industrial Statistics (ISBIS), European Network for Business and Industrial Statistics (ENBIS).

We hope this volume will serve as a source of inspiration and a valuable reference for researchers working at the interface of statistics, data science, and the evaluation of quality in contemporary society.

Marica Manisera

Chair of the Program Committee

Giovanna Boccuzzo

Co-chair of Local Organizing Committee

Luigi Salmaso

Co-chair of Local Organizing Committee

Maurizio Carpita

Coordinator of the SVQS Group

Contents

Plenary Sessions	1
Unsupervised Statistical Tools for Anomaly Detection: The Case of Healthcare Frauds <i>F. Ruggeri</i>	2
New data science breakthroughs in sports (medicine) <i>C. Ley</i>	2
A unifying approach to the Behrens-Fisher problem <i>F. Pesarin</i>	3
Recent challenges in optimal design of experiments (dedicated to the memory of H.P. Wynn) <i>W.G. Müller</i>	4
Invited session IPS01 - CLADAG - <i>Advances in statistical modeling of educational data</i>	5
Combining multiple imputation and variable selection for the analysis of Invalsi test data <i>S. Bacci, E. Dreassi, L. Grilli, and C. Rampichini</i>	6
A latent class model for item responses and response times <i>D. Spinelli, F. Pennoni, F. Bartolucci, and G. Vittadini</i>	13
Exploring factors affecting gender gap in university student performance <i>M. La Rocca, M. Niglio, M. Restaino, and M.P. Vitale</i>	19
Invited session IPS02 - SIS-ySIS - <i>Advanced Methods for Assessing Environmental Impact and Green Innovation</i>	27
Clustering Hierarchical Disjoint Principal Component Analysis for Environmental Impact Assessment of Food Products <i>M. Bottazzi Schenone, M. Vichi</i>	28
Spatiotemporal clustering of GHGs emissions in Europe: exploring the role of spatial component <i>C. Morelli, P. Maranzano, and P. Otto</i>	36
Invited session IPS03 - <i>Evaluation in transportation and mobility: different points of view</i>	45

Greenhouse gas emissions and audit fees in the transportation and logistic sector <i>R. Santonastaso, R. Macchioni, and A. D'Ambra</i>	46
The environmental assessment of sustainable vehicular transport through the Life Cycle Analysis (LCA) methodology applied to a vehicle <i>L. Della Ragione, and G. Meccariello</i>	54
Towards a more accurate prediction of aircraft passenger discomfort by integrating temporal analysis of seat-interface pressure with static passenger anthropometric features <i>A. Vanacore, and A. Ciardiello</i>	62
Invited session IPS04 - Functional data analysis in action	71
Fuzzy clustering of NPL recovery curves <i>A. Carleo, and R. Rocci</i>	72
Enhancing survival learning from time-to-event data with functional random survival forest <i>G. Loffredo, E. Romano, and F. Maturo</i>	79
Leveraging Functional Diversity in Statistical Ensemble Learning for Robust Functional Data Classification <i>D. Riccio, F. Maturo, and E. Romano</i>	86
Forecasting carbon emissions from electricity generation: classical vs functional methods <i>P. Dutillo, and F. Lisi</i>	93
Invited session IPS05 - SIS-SDS - Statistics and data science for sustainable science	101
Learning from Crowds: Model-Based Approaches with a Novel Discriminant Analysis Method <i>A. Cappozzo</i>	102
Multi-state Markov-switching for extreme summer temperatures modelling <i>G. Di Credico, V. Gioia, and F. Pauli</i>	108
Bayesian Modelling for Forensic Evaluation of Handwritten Loop Characters <i>L. Tzai, I. Ntzoufras, F. Taroni, and S. Bozza</i>	116
A copula-based hidden semi-Markov model for toroidal data <i>M. Mingione, and F. Lagona</i>	123
Invited session IPS06 - Frailty in Aging Populations	131
Measuring frailty in the elderly: an indicator based on a combination of classifiers <i>P. Belloni, and S. Rebottini</i>	132
Where to age happily: a quantitative assessment of Age-Friendliness in the districts of three Italian cities <i>L. Schiavon, and M. Stival</i>	140

Social Frailty and Healthcare Utilization Among Community-Dwelling Older Adults in Europe <i>E. Cisotto, A. Donno, G. Cavrini, and G. Boccuzzo</i>	147
Invited session IPS07 - Well being: humans and algorithms	155
Assessing the impact of Active and Assisted Living technologies on older adults’ psychosocial well-being: there is room for improvement <i>A. Bodini, A. Brancaccio, C.G. Leo, P. Mincarone, V. Recchia, and M.R. Tumolo</i>	156
RAISE Math: An Adaptive Approach to Middle School Mathematical Proficiency <i>A. Brancaccio, and A. Bodini</i>	163
Thermal Comfort Monitoring of Urban Areas: A Spatio-Temporal Jump Model Approach <i>F.P. Cortese, and A. Pievatolo</i>	170
Assessing the SAFETy of Machine Learning Models <i>E. Raffinetti, and A. Spelta</i>	178
Invited session IPS08 - Advancing Social Sustainability: Approaches and Frameworks	187
The CARESWELL Project on professional well-being in elderly care <i>L. Salaris, N. Tedesco, M. Abate, G.M. Milanese, A. Spera, and G. Tricarico</i>	188
Tourism and Economic Growth in Basilicata: The Impact of Matera’s European Capital of Culture <i>M.G. Grassia, M. Marino, R. Cataldo, S. Paesano, R. Mazza, D. Sacco, and A. Stavolo</i>	196
SustVIS: bridging Corporate Reporting and actionable sustainability insights <i>A. Forciniti, and E. Zavarrone</i>	204
Invited session IPS09 - Interplay between multimorbidity and frailty: implications for health and aging	211
Sensitivity Intervals for a Frailty Indicator based on partially ordered set theory <i>M. Nicolaio, M. Silan, and G. Boccuzzo</i>	212
On longitudinal study of chronic diseases multimorbidity patterns at older ages <i>L.H.K. Dang, N. Caranci, G. Roli, R. Rettaroli, and R. Miglio</i>	220
Measuring frailty through the multimorbidity network <i>E. Banzato, and G. Boccuzzo</i>	228
Invited session IPS10 - Assessing Young Students’ Academic Performance: Insights from Recent Studies on Digital and Mathematical Competence	237
Digitalization and Digital Competences. Towards a Measurement Model at the End of Compulsory Schooling <i>S. Ercolanoni, and M. Gui</i>	238

Investigating inequalities in digital skills of Italian students based on ICILS data <i>L. Bungaro, M. Matteucci, and S. Mignani</i>	245
Measuring Digital Competencies in Secondary Education: A First-Phase Investigation of Standardized Assessment and Certification Models <i>P. Falzetti, P. Giangiacomo, and R. Ricci</i>	252
Evaluating Student Performance: Unveiling the Impact of School's Economic, Social, and Cultural Status with Regression Discontinuity <i>P. Sannino, C. Davino, and R. Romano</i>	258
Invited session IPS11 - Advances in Functional Data Analysis and Applications	267
Multivariate functional data analysis to monitor, evaluate and classify engineering failures <i>C. Yildirim, A.M. Franco-Pereira, and R.E. Lillo</i>	268
Partially observed functional data: a general model <i>M.C. Aguilera-Morillo, M. Durban, and P. Hernández-Amaro</i>	276
Basis expansion approaches for penalized functional ICA with applications in neuroscience <i>A.M. Aguilera, and M. Vidal</i>	281
New Perspectives on Explainability Tools for Functional Random Forests <i>F. Mauro, and A. Porreca</i>	288
Invited session IPS12 - Futures Studies and Statistics	295
Policy analysis with generative pre-trained transformers: the future of north-eastern Italian families <i>Y. Calleo, M. Scioni, S. Di Zio, M. Marozzi, and M. Bolzan</i>	296
A classification of Delphi outputs in scenario development adopting a network-based fuzzy clustering <i>Y. Calleo, M. Bolzan, and S. Di Zio</i>	303
Predicting and Preventing gender-based violence: A strategic framework for long-term change <i>E. Zavarrone, and E. Forciniti</i>	311
A spatio-temporal forecasting model for the State of the Future Index of European countries <i>L.S. Alaimo, Y. Calleo, and S. Di Zio</i>	319
Invited session IPS13 - Digital Statistics and Official Data: Innovation, Quality, and Challenges	327
Ensuring Data Quality in Population and Housing Census in Albania: Challenges and Solutions in the Digital Transformation Era <i>E. Dhuli, and K. Sevrani</i>	328

Household income and academic outcome: A register-based analysis of Italian students <i>L. Coppola, D. Ichim, and I. Siciliani</i>	334
Improving official statistics: the new Audit internal procedure at Istat. Description and findings <i>M.F. Loporcaro</i>	341
Introducing AI in Istat: Responsibility and Innovation <i>S. Letardi, and F. Orecchini</i>	348
Invited session IPS14 - Advances in Statistical Learning from High-Dimensional Data	355
Different dimension reduction approaches for estimating the functional regression model <i>H. Ortiz Alcalà, C. Acal, M. Vidal, and A.M. Aguilera</i>	356
Augmented Functional Random Forests: Robust Classifier Construction and Conditional Importance of Functional Principal Components <i>F. Mauro, and A. Porreca</i>	363
Functional Clustering for Survival Curves <i>M. De Lucia, E. Romano, and F. Mauro</i>	370
Invited session IPS15 - BDsports - Sport Statistics	379
Fitting Bayesian and frequentist football models in R via the footBayes package <i>L. Egidi, and R. Macrì Demartino</i>	380
Improved prediction of 100-meter sprint records <i>G. Fonseca, F. Giummolè, M. Lambardi di San Miniato, and V. Mameli</i> . . .	387
The role of basketball shooting trajectories in predicting shot success <i>A. Macis, M. Sandri, and P. Zuccolotto</i>	393
Exploring the Role of Kinematic Variables in Determining Match Outcomes in Serie A: A Statistical Approach <i>A. Lucadamo, P. Troiani, F. Laterza, C. Savoia, M. Bertollo, D. Pompa, S. Pacillo, P. Amenta, and L. Greco</i>	401
Invited session IPS16 - ISBIS - Statistical Modeling for New Economic and Social Insights	409
Exploring Spatial Disparities and Gaps in Transportation and Mobility through Airbnb Reviews: A Sardinian Case Study <i>L. Frigau, G. Contu, M. Ortu, and F. Mola</i>	410
Web-Scraped Data for MSMEs Economic Analysis: Insights on Innovation and Bankruptcy <i>C. Bottai, L. Crosato, and C. Liberati</i>	418
Invited session IPS17 - Sustainable tourism for territorial development	427

Multidimensional Statistical Indicators For The Analysis Of Sustainable Tourism <i>P. Perchinunno, A. Massari, S. L'Abbate, A.M. D'Ugento, and C. Crocetta</i> .	428
Tourism and territories: a multilevel analysis in Italy <i>L.S. Alaimo, E. di Bella, and S. L'Abbate</i>	435
Sustainable Tourism Indicators: A Classification Proposal from the SDGs Perspective <i>V. Tomaselli and M.D. Giammanco</i>	442
Invited session IPS18 - New methods and applications for the analysis of survey data	449
A distance-based consensus measure for analyzing ternary preferences <i>A. Albano, J.L. Garcia-Lapresta, A. Plaia, and M. Sciandra</i>	450
Lost in Noise: When cleaning up clouds the picture. Fuzzy topic modeling and robust low-rank decomposition <i>A. Calcagni, A. Sciandra, and A. Tuzzi</i>	458
Assessing Latent Structure in Questionnaire Data: an integrated approach between Consensus Bayesian Networks and Community detection <i>P. D'Urso, L. DeGiovanni, L. Federico, and V. Vitale</i>	465
Composite Likelihood Inference for Simultaneous Clustering and Dimensionality Reduction in the Chinese Longitudinal Healthy Longevity Survey <i>M. Ranalli, R. Rocci, and A. Maruotti</i>	473
Invited session IPS19 - Luxembourg Statistical Society - Statistically Enhanced Learning	479
Estimation of match abilities for tennis players via a maximum likelihood approach <i>H. Bartmann, A. Groll, and R. Michels</i>	480
Variable ranking and data reduction in GLM domain <i>F. Giordano, M. Niglio, and M. Restaino</i>	488
Invited session IPS20 - Advances in multivariate permutation methods	495
Advances in Cluster Analysis: a permutation approach to cluster validation <i>E. Barzizza, N. Biasetton, and R. Ceccato</i>	496
A Nonparametric Approach for Robust Product Ranking <i>E. Barzizza, and R. Ceccato</i>	503
Comparing Multiple Samples under Heteroscedasticity: a Multivariate Permutation Solution <i>E. Barzizza, N. Biasetton, and R. Ceccato</i>	510
Nonparametric Time Series Screening for Bitcoin and Gold Volatility <i>G. Feo, F. Giordano, and S. Milito</i>	518
Invited session IPS21 - Measuring latent traits and perceptions for social innovation	527
Streaming entropy for dynamic modelling in psychological measurement <i>M. Angelelli, O.C.G. Gelo, and E. Ciavolino</i>	528

Low-Rank Analysis of Topic Quality: Comparing LDA, CTM, and Fuzzy-LSA methods <i>A. Calcagni</i>	535
A Penalized Complexity Prior for the Measurement Error of Covariates in Small Area Models <i>D. Battagliese, A. Pollice, S. Arima, and B. Liseo</i>	543
Patient satisfaction and healthcare quality perception: an explorative textual analysis <i>C. Valentino, M. Aria, M. Angelelli, and E. Ciavolino</i>	551
Invited session IPS22 - Advances in the Modelling of Spatio-Temporal Data	559
Transition Models for Precipitation Climatology Estimation <i>N. Umlauf, and R. Stauffer</i>	560
Spatial and Spatiotemporal Volatility Models: A Review <i>P. Otto, O. Doğan, S. Taşpınar, W. Schmid, and A.K. Bera</i>	568
Signature Lasso for Fractional Brownian Motion Processes <i>A. Mohaddes, F. Iafrate, and J. Lederer</i>	575
Invited session IPS23 - Reporting violence against women: statistical tools and some pitfalls	583
A Territorial Analysis of Factors Affecting Access to Support Centers for Violence Against Women in Apulia <i>C. Calculli, L. Cefalo, and A. Pollice</i>	584
A Bayesian dynamic model for the analysis of 1522 helpline data <i>S. Polettini, S. Arima, and A. Farcomeni</i>	591
Under-reported data: a simulation study about parameter sensitivity <i>G. Panunzi, S. Polettini, and S. Arima</i>	599
Invited session IPS24 - SIS-DIAS - Data science to measure environmental sustainability	607
Working with composite indicators to measure environmental sustainability <i>M. Ciommi, F. Mariani, G. Polinesi, and M.C. Recchioni</i>	608
Assessing CO2 emissions from electricity generation: a methodological review and comparative analysis <i>M. Bertolini, P. Dutillo, and F. Lisi</i>	616
Invited session IPS25 - BDsports - Sport statistics: methods and applications	623
A New Dataset for Exploring Actions of Italian Football Matches <i>M. Cefis, R. Metulini, and M. Carpita</i>	624
A Financial Instrument Indexed to Football Players Performance <i>G. Dello Iorio, and M. Cefis</i>	631

Towards Smarter Player Scouting: Learning Football Player Embeddings with Variational Autoencoders (VAEs) <i>G. Fantuzzi, L. Egidi, and N. Torelli</i>	637
On the relationship between passing-related features and the result of offensive actions in football <i>R. Ievoli, L. Palazzo, R. Rondinelli, and G. Ragozini</i>	645
Invited session IPS26 - Advanced methods in Labour Market	653
Dynamics of IT and Digital Professional Profiles in Italy: An analysis of the trajectories of required soft skills <i>P. Mariani, A. Marletta, D. Pirotta, and M. Zenga</i>	654
Hospitality and Tourism Activities in the Italian Provincial Labour Market <i>P. Mariani, and M. Zenga</i>	661
Gender Diversity in Life Sciences Firms: A Data Integration Approach <i>L. Benedan, C. Colapinto, P. Mariani, L. Pagani, and M. Zenga</i>	668
Gender inequality in the workplace: a comparative analysis of Italian companies through the definition of composite indicators <i>A. Gattulli, M. Antonicelli, I. De Turi, and M. Posa</i>	674
Invited session IPS27 - New methods for university assessment in teaching and research	683
Research assessment in Italy: the recent experience and the way forward <i>M. Malgarini, and S. Sarlo</i>	684
Insights into the Italian university system: analyzing standardized indicators <i>I. Camminatiello, and R. Lombardo</i>	691
Can the dimension of a university affect the assessment of the performance? <i>A.M. D'Ugento, V. Ricci, and E. Toma</i>	698
Invited session IPS28 - ENBIS - Handling uncertainty and complexity in modern decision-making and smart applications	707
A Bayesian model for measurements by counting <i>F. Pennecchi, and W. Bich</i>	708
The benefits of classification: a system operational perspective <i>Y.N. Marmor</i>	716
A Scalable and Cost-Effective Approach for Disease Detection in Smart Agriculture: the case of Xylella Fastidiosa <i>A. Vanacore, A. Ciardiello, L. Uccello, G. P. Auricchio, and A. Izzo</i>	723
Invited session IPS30 - High dimensional data analysis for evaluating sustainable development	731
Identity Spline and LASSO Regression for Analysing Italian Poverty <i>R. Lombardo, M. Musella, and I. Camminatiello</i>	732

Evaluation of Synthetic Indicators: a new method based on functional data analysis <i>L.S. Alaimo, F. Fortuna, and A. Naccarato</i>	739
From Data Collection to Knowledge in Health: Turning Scientific Production into Practical Decision-Making Tools <i>M. Aria, C. Cuccurullo, L. D’Aniello, M. Spano, and C. Alabiso</i>	746
Invited session IPS31 - <i>Statistical methods for measuring the sustainability</i>	755
Toward a Sustainable Society: assessing legal progress of LGBT rights with func- tional data analysis <i>F. Fortuna, A. Naccarato, and A. Trezzini</i>	756
Advanced multidimensional spatial clustering for sustainable transport assessment <i>V. Distefano and S. Deiacò</i>	763
Quantile-Based Hidden Markov Graphical Models for PM10 Pollution Analysis: A Perspective on Air Quality in Northern Italy <i>B. Foroni, L. Merlo, L. Petrella, and N. Salvati</i>	770
Invited session IPS32 - Austrian Statistical Society - <i>Effectively Using Relevant Infor- mation for Statistical Inference</i>	779
How to efficiently analyse data in rare diseases research: Accounting for individual variability of the primary outcome over time <i>G. Zimmermann</i>	780
Random Reward Phase-Type Distributions with Applications in Employee Devel- opment <i>S. Pauli, and A. Futschik</i>	787
How to best analyze randomized experiments with very small samples <i>S.L. Grüniger, R. Furrer, and F. Frommlet</i>	793
Nonparametric Analysis of Tourism Data for Optimizing Sustainable Transport Services at Airport Salzburg <i>M. Kiefel, and D. Lahnsteiner</i>	800
Invited session IPS33 - <i>Statistics and geometry with applications</i>	809
An improved TDA-based algorithm for the isomorphism problem of orthogonal arrays <i>R. Fontana, and M. Guerra</i>	810
A geometrical approach to research assessment <i>B. Cappelletti-Montano, S. Columbu, S. Montaldo, and M. Musio</i>	818
Impact of zero-patterns in the dependence of three-way binary tables <i>R. Fontana, E. Perrone, and F. Rapallo</i>	826
Three dimensional Bernoulli generators and their dependence structure <i>R. Fontana, and P. Semeraro</i>	833
Invited session IPS34 - <i>Nonparametric methods and applications</i>	841

Statistical Analysis of the Italian pension system: a nonparametric approach <i>M. Giacalone, G. Piscopo, S. Marra, and C. Marciano</i>	842
Impact assessment of big music events with a focus on the food service industry: a nonparametric test for linear models <i>M. Borghesi</i>	849
Local Inference for Data Giants <i>G. Cattani, M. Scholz, and S. Sperlich</i>	856
Time series comparisons with application to electricity market <i>S. Bonnini, and M. Borghesi</i>	863
Invited session IPS35 - <i>Statistics and AI to support sustainability in agro-ecosystems</i>	871
Quantitative Tools for Evaluating Biological Processes in Vineyard Agroecosystems: BIOPASS [®] and QBS-ar <i>M. Tonni, P. Donna, M. Donna, I. Ghiglieno, L. Valenti, and G. Gilioli</i> . . .	872
Impact of meteorological conditions, canopy shading and leaf removal on yield, must quality, and norisoprenoid compounds content in Franciacorta sparkling wine <i>S. Carlin, I. Ghiglieno, G. Cola, U. Vrhousek, L. Valenti, M. Garcia-Aloy, and F. Mattivi</i>	880
Leveraging Remote Sensing and Artificial Intelligence to Monitor Biodiversity in Wine-sector and Semi-Natural Urban Environments <i>G. Tariku, I. Ghiglieno, A. Simonetto, I. Serina, and G. Gilioli</i>	887
Measuring supply chain sustainability: a structured multilevel framework <i>F. Demaria, F. Bertacchini, U. Kocollari, and M. Cavicchioli</i>	894
Invited session IPS36 - <i>Business Intelligence for Official Statistics: Tools for Improving Quality</i>	903
Artificial intelligence and methodological innovations in the statistical process: a GSBPM perspective <i>G. Ascari</i>	904
Business Intelligence for assessing Istat HR processes <i>A. Dentini, and M. Mazziotta</i>	912
The Importance of Business Intelligence in Monitoring Organizational Effectiveness: Analysis of some organizational structures in Istat <i>S. Demofonti, and S. Pifferi</i>	920
Statistical indicators on public administration employees: what support from Business Intelligence? <i>G. Cinquegrana, R. Fantozzi, and I. Screpante</i>	928
Invited session IPS37 - <i>Statistical methods for traffic and the environment</i>	935

Assessing Predictive Performance of Time Series Copula Models in Traffic Flow Analysis <i>S.S. Guerini, and R. Metulini</i>	936
Traffic Restrictions and Air Quality: A Counterfactual Matrix Completion Analysis of Milan’s Area C <i>R. Adam, F. Biancalani, and R. Metulini</i>	944
Characteristic-based clustering of traffic flow time series with complex seasonality <i>M. Carpita, R. Metulini, and M. Migliorati</i>	952
Graph Neural Networks for Traffic Accident Risk <i>G. Cantaluppi, J. Lin, G. Valena, and D. Zappa</i>	959
Invited session IPS39 - MICS - Data-Driven Strategy, Management, and Technical Development for ‘Made in Italy’ Products	967
The VABSA framework: a Value- and Attribute-Based Sentiment Analysis of consumer perceptions <i>N. Biassetton, R. Ricciardi, and N. Arghistani</i>	968
Group-specific document embeddings: An application to User-Generated Content <i>F. Alberti, R. Ricciardi, and M. Manisera</i>	975
Model-Based Clustering of Multivariate Rating Data to Evaluate Consumer Perceptions on Sustainability <i>M. Ventura, J. Jacques, and P. Zuccolotto</i>	983
Permutation-Based Multivariate Chi-Square Test: A Case Study <i>M. Disegna, A. Fanesi, and L. Salmaso</i>	991
Invited session IPS40 - Advancing Sustainability in Healthcare	999
Early dementia recognition with Digital Linguistic Biomarker <i>L. Altieri, and G. Gagliardi</i>	1000
Inequalities in adherence to the diabetes guidelines in the migrant population at the regional level <i>M. Arcaio, A. Allotta, S. Scondotto, and M. Ferrante</i>	1007
Research excellence and patient perception: investigating the impact of AHSCs’ scientific output <i>A. Gnasso, D. Sacco, L. Celardo, M.A. Smecca, C. Alabiso, and M. Spano</i>	1014
Invited session IPS41 - Multivariate tools for Sustainability	1021
Identifying optimal candidate models of classification trees with weak bias <i>K. Hayashi, and R. Inose</i>	1022
A Synthetic Green Indicator and Its Role in Green Bond Pricing <i>P. Galfrascoli, E. Ossola, and G.S. Monti</i>	1029
Sustainability in sports: a three-decade scientific review <i>L. Palazzo, T. Menini, V. Simonacci, and M. Gallo</i>	1036

Do SMEs really walk the talk? A textual analysis of sustainability reporting in Italy <i>A. Sarra, B. Iannone, M. Di Giampietro, A. Evangelista, and T. Di Battista</i>	1043
Invited session IPS42 - <i>Tourism and sustainability</i>	1051
New approaches for innovative tourism promotion among artificial intelligence and the metaverse <i>V. Santarcangelo, M. Giacalone, A. Romano, and G. Piscopo</i>	1052
A proposal for daily mortality indices related to the official measurement of pollutants <i>S. Cervellera, M. Giacalone, M. Longobardi, G. Piscopo, and C. Cusatelli</i>	1060
An ultrametric composite indicator for ranking the environmental sustainability performance of Italian provinces <i>I. Ariante and P. Sarnacchiaro</i>	1067
Invited session IPS43 - <i>Data Analytics for Mobility Domains</i>	1073
Quality Key Factors in the Public Transportation Scenario <i>S. Porcu, and F. Marcello</i>	1074
A review and future directions in big data subsampling <i>R. Arboretti and A. Molena</i>	1082
Regulating Spatial Fairness in a Tripartite Micromobility Sharing System via Reinforcement Learning <i>M. Cederle, M. Fabris, and G.A. Susto</i>	1089
Overcoming Computational Challenges in Two-Stage Multi-path Traveling Salesman Problem via Neural Networks <i>X. Chou, L. di Marco, and E. Messina</i>	1097
Contributed Session CPS01 - <i>Statistical methods and models for student mobility, performance, and resilience</i>	1105
Dropout rates in Italian universities: a comparison of online and traditional universities <i>C. Usala, G. Lombardi, and A. Priulla</i>	1106
Navigating Academic Waters: Exploring the Trends of Switching Students in Higher Education <i>V. Tocchioni, S. Milone, G. Lombardi, and A. Petrucci</i>	1114
Evaluating the Impact of Teachers Characteristics and Practices on High School Students' Performance <i>I. Moreschini, N. Brancaccio, I. Sulis, and M.P. Vitale</i>	1121
Optimizing goodness of fit of mixture models with uncertainty for rating data: a case study on academic distress <i>R. Simone</i>	1129

Bayesian Downscaling of INVALSI Data with Small Area Estimation <i>D. Battagliese, and A. Pollice</i>	1137
Academic Resilience Among Low ESCS Students in Italy <i>A. Pisu, I. Sulis, and M. Porcu</i>	1145
Contributed Session CPS02 - Statistical modeling for economic and social studies	1153
Assessment of the diversity of housing expenditures of households in Polish regions <i>M. Sadko, and B. Pawelek</i>	1154
Linking Income and Housing Poverty: A Comprehensive Analysis Using Integrated Fuzzy and Relative Approach <i>M. Ćwiek, P. Ulman, and M. Sadko</i>	1161
Disability and Economic Disparities in Europe: A 2023 Comparative Analysis <i>P. Ulman, and A. Wałęga</i>	1168
Social Protection and Household Food Security: Evidence from Ethiopia’s PSNP Program <i>S.T. Mengesha, E. Stanghellini, and M. Genbäck</i>	1175
Supportive Obligations among Nuclear and Extended Kin: A Vignette Study <i>M.F. Morabito, M. Tosi, and V. Tocchioni</i>	1182
Material deprivation in Europe: a comparative analysis of the phenomenon through demographic parameters <i>M. Antonicelli</i>	1189
Contributed Session CPS03 - Statistical modeling for environmental, social, and economic challenges	1197
Model confidence set for air pollution time series models <i>G. Zammarchi, and Y. Kawasaki</i>	1198
Risk map for female genital mutilation/cutting <i>S. Golia, D. Carrillo, P. Farina, L.E. Ortensi, and E. Ripamonti</i>	1205
Is Time Well Spent in Later Life? Understanding Life Satisfaction Through Daily Time Reallocation Across Activities <i>M. Salierno, A. Donno, and M.L. Tanturri</i>	1212
Adaptive Bayesian Spatial Modeling of Economic Resilience to Environmental and Climate Shocks <i>A. Pacifico</i>	1220
Exploring Peer Effect in Educational Choices: The Missing Tie <i>A. Pacca, and G. Giordano</i>	1228
If I say bread... What comes to your mind? <i>S. Nascetti, D. Doglio, F. Ciesa, D. Delvai, P. Selle</i>	1235
Contributed Session CPS04 - Advances in statistical methodology and computational techniques	1243

A comparative analysis of machine learning algorithms for predicting probabilities of default <i>A.I. Cristescu, and M. Giordano</i>	1244
Segmenting the spatial distribution of the adjusted dissimilarity index to detect residential segregation of foreigners in Campania <i>R. Simone, and F. Benassi</i>	1251
Extreme-value copulas: a note on comparing models for tail-risk events <i>A. Gatto, S. Perazzini, F.M.L. Di Lascio</i>	1258
On a generalized stochastic Bass model <i>G. Albano, A. Barrera, V. Giorno, and F. Torres Ruiz</i>	1266
On CTA-PLS corrections applied on sports performance <i>M. Angelelli, M. Cefis, E. Ciavolino, and Maurizio Carpita</i>	1273
Skew-normal finite mixture models for change-point detection in time series <i>M. Forti, A. Nigri, and H. Shang</i>	1280
Contributed Session CPS05 - Statistical and computational methods for mobility, tourism, and risk analysis	1287
Generative AI for traffic scenarios: a GCN-VAE model <i>M. Carbonera, M. Ciavotta, and E. Messina</i>	1288
A Markov chain approach to investigate the Tourists movements: the case of Sardinia <i>G. Contu, L. Frigau, M. Ortu, and F. Mola</i>	1296
Identifying corruption risk profiles in public procurement over emergency periods: a latent class - Item Response Theory approach <i>S. Del Sarto, M. Gnaldi, and N. Salvini</i>	1303
Classification and Forecasting of Foreign Tourists Using Time Series from Mobile Phone Networks <i>R. Rossetti, and Y. Sagaert</i>	1310
Denosing and Forecasting Model for Structural Health Monitoring of Bridges: A Case Study in Brescia <i>L. Biazzo, and L. Fenga</i>	1318
The cost of isolation: how insularity affects European transport cost <i>G. Zammarchi and P. Mattana</i>	1326
Contributed Session CPS06 - Advances in text mining and visualization of complex data	1333
A Novel Metric for Enhancing Online Review Relevance in E-commerce <i>L. Stracqualursi, and P. Agati</i>	1334
The APP-LIC-IT data visualization tool <i>G. Cantaluppi, D. Zappa, M. Rho, and J. Vera Bacigalupo</i>	1341

CONTENTS

From Self-Perception to Feedback: Mapping Sustainability-Oriented Self-Descriptions to Airbnb Reviews <i>M. Ortu, G. Contu, and L. Frigau</i>	1349
Natural Language Processing for Data Analysis: an application on the well-being <i>S. Boi, N. Tedesco, M. Romano, and L. Salaris</i>	1356
Constructing fair and comprehensive dataset of online opinions for human annotations: considerations and challenges dealing with Italian content <i>A. Cucco, E. del Gobbo, L. Fontanella, S. Fontanella, and L. Ippoliti</i>	1363
The Impact of Inappropriate Emergency Admissions in the Autonomous Province of Bolzano: Challenges and Insights <i>M. Bonetti, G. D’Orio, and C. Seghieri</i>	1370

Invited session IPS24 - SIS-DIAS - *Data science to measure environmental sustainability*

Organiser: Francesca Bassi

Chair: Francesca Bassi

Discussant: Leonardo Alaimo

- M. Ciommi, F. Mariani, G. Polinesi, and M.C. Recchioni - Working with composite indicators to measure environmental sustainability
- M. Bertolini, P. Dutillo, and F. Lisi - Assessing CO2 emissions from electricity generation: a methodological review and comparative analysis

Working with composite indicators to measure
environmental sustainability

Mariateresa Ciommi¹, Francesca Mariani¹,
Gloria Polinesi¹, Maria Cristina Recchioni¹

¹Università Politecnica delle Marche

Working with composite indicators to measure environmental sustainability

Mariateresa Ciommi, Francesca Mariani, Gloria Polinesi, Maria Cristina Recchioni

Abstract The construction of composite indicators often involves the aggregation of several indicators into a single measure. These individual indicators typically represent different aspects of the phenomenon being investigated, such as economic, social, or environmental factors. Since these indicators are measured in different units, scales, and ranges, a critical step in the construction is the normalization step. Normalization transforms data into a common scale enabling meaningful comparison and allowing for aggregation. However, normalization introduces several challenges and potential pitfalls that can affect the validity and interpretability of composite indicators. Here, we focus on the max-min normalization method, and we modify it by allowing to vary the classical minimum and maximum values in a given interval, respectively. Adjusting these values within a range generates a distribution of normalized values for each indicator. Consequently, we can take the mean value and the corresponding error of measurement as normalized value of the indicator. The error can be interpreted as a sort of degree of reliability associated with the mean value. Finally, we aggregate those values along with their errors of measurement. An empirical application, based on environmental data for Italian Provinces, illustrates how the method works.

Keywords: Composite indicators, Normalization, max-min method, environmental sustainability.

1. Introduction

In the construction of a *composite indicator*, researchers have to face a common practical problem, namely how to make comparable indicators. This is the well-known normalization step. While on the one hand, many articles have been written about the construction of indicators, the preliminary normalization step has received less attention. However, the choice of a normalization method is not harmless since different methods yield different results, which can influence outcomes ([1]). Literature accounts for several normalization methods, such as: i) *Ranking of indicators across countries*¹; ii) *Standardization (or z-*

¹This is the simplest normalization approach consisting in ranking each indicator across countries. Despite its simplicity and resilience to outliers, the method has a significant drawback: it results in the loss of information about absolute levels and makes it impossible to draw conclusions about differences in performance.

scores)²; iii) *Re-scaling*³; iv) *Distance to a reference country*⁴; v) *Categorical scales*⁵; vi) Indicators above or below the mean⁶ (see [2] for a detailed review on normalization methods.).

The approach we propose defines as normalized value for a given unit the mean of the distribution of normalized indicators with a measurement error that represent the uncertainty around this value. In this way, each unit is characterized by a value (the mean) and its own confidence interval, and two units are considered significantly different if their intervals do not overlap.

The method is used to represent uncertainty or variability in the data, providing a range of possible values rather than a single point estimate (see, for example [3]). Intervals for composite indicators also appear in [4] but in that paper authors fix a normalization and then use a Monte Carlo simulation to account for several possible systems of weights in the aggregation step.

The method we propose has several advantages. Firstly, it introduces robustness into the ranking since two units are really different if their intervals do not intersect. Secondly, the expansion of the possible ranges for the minimum and maximum value allows for temporal comparisons. In fact, with this approach, when a new year is added, it may not be necessary to perform a new normalization since the new values of maximum and minimum could belong to the ranges mentioned above. Finally, from a technical point of view, the normalized values fall in the range (0, 1) and, since it avoids zero values, in the aggregation phase we can apply any generalized mean [5], which in general only works for strictly positive values.

In recent years, measuring environmental sustainability has become increasingly important due to the growing challenges regarding climate change, biodiversity loss, and resource

²Normalized values are obtained by subtracting the average value across countries from the raw value and dividing by the standard deviation. This standardization is widely used because it converts all indicators to a common scale (average = 0, standard deviation = 1), preventing aggregation distortions caused by differing means.

³Normalized values are calculated as the ratio of the difference between the raw value and the minimum value, divided by the range, as consequence, normalized values range in [0, 1]. It is also known as *max-min method*. This approach is sensitive to unreliable outliers, as minima and maxima can distort results. It also amplifies the influence of indicators with small value intervals on the composite indicator. For time-dependent studies, the range is often fixed across the entire time frame. If new data exceed the selected range, the composite indicator must be recalculated for all years to ensure comparability.

⁴This method normalizes an indicator by dividing its value for a given country at a specific point in time by the value of a reference country at an initial time. This approach accounts for the evolution of indicators over time.

⁵To get normalized values, each indicator is assigned a categorical score, which may be numerical or qualitative (e.g., "fully achieved," "partly achieved," "not achieved"). In some cases, the scores are determined based on the percentiles of the indicator's distribution across countries. Categorical scales offer the advantage that small changes in the indicator value (e.g., over time) do not impact the normalized value. However, this can also be a drawback, as a significant amount of information about the variance between countries in the normalized indicators is lost.

⁶This normalization method differentiates between values that are above, near, or below an arbitrarily defined percentage threshold around the mean. The normalized value is set to 1 if the indicator exceeds the threshold, -1 if it falls below, and 0 if it lies within the neutral zone around the mean. The method's simplicity and robustness against outliers are its key advantages, while its drawbacks include the arbitrary nature of the threshold and the loss of information about absolute levels.

depletion. Thereby, measuring environmental sustainability at local level, such as Regions (NUTS2) or Provinces (NUTS3) is crucial for tailoring local policies for specific ecological needs and understanding of regional differences in air quality, water usage, energy consumption, and waste management. Focusing on provincial-level data allows policymakers to design more effective interventions than national-level strategies while ensuring better resource allocation.

Here, we apply our approach to measure environmental sustainability of the Italian Provinces using the *Sole24ore* data collected in 2024.

The rest of the paper is organized as follows. Section 2 introduces the method and Section 3 illustrates its potential by analyzing environmental data. Section 4 concludes.

2. The method

Let X be a $n \times k$ matrix whose element x_{ij} represents the value of the j -th elementary indicator $j = 1, \dots, k$ for the i -th local unit (e.g., the Italian province) $i = 1, \dots, n$. We denote by r_{ij} the normalized value, obtained according to the classical *min-max method*:

$$r_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}}. \quad (1)$$

where, $\min_i x_{ij}$ and $\max_i x_{ij}$ denote the minimum and maximum value across all units, respectively

As stressed in Section 1, the choice of maximum and minimum can lead to different results. Moreover, adding a new year may require to update values from previous years. For this reason, we modify the method as follows: we fix a *maximum*, an *upper bound*, UB and a *minimum*, a *lower bound*, LB external to the intervals and proportional to the real $\max_i x_{ij}$ and $\min_i x_{ij}$ for each variable j .

Then, we normalize the data by choosing as maximum and minimum value in (1) any possible number in the interval $[\max_i x_{ij}, UB(x_{ij})]$ and $[LB(x_{ij}), \min_i x_{ij}]$, respectively. More in details, for the maximum (minimum), we move from $\max_i x_{ij}$ ($LB(x_{ij})$) to UB (LB) by ϵ_U (ϵ_L) > 0 , that is

$$\epsilon_U = \frac{UB(x_{ij}) - \max_i x_{ij}}{\delta_U} \quad \epsilon_L = \frac{\min_i x_{ij} - LB(x_{ij})}{\delta_L} \quad (2)$$

where δ_L, δ_U are the number of parts into which we want to divide the interval and without loss of generality, we can fix $\delta = \delta_L = \delta_U$, with $\delta = 100$. in our empirical analysis. Consequently, equation (1) can be re-written as follows:

$$Z_{ij}^v = \frac{x_{ij} - a_v}{b_v - a_v} = \frac{x_{ij} - (LB + v \cdot \epsilon_L)}{(UB - v \cdot \epsilon_U) - (LB + v \cdot \epsilon_L)} = \frac{x_{ij} - LB - v \cdot \epsilon_L}{(UB - LB) - v \cdot (\epsilon_U + \epsilon_L)} \quad v = 0, \dots, \delta \quad (3)$$

As a consequence, for each v , $v = 0, \dots, \delta$ we get a (column-)vector of normalized values starting from x_{ij} and we collect these vectors in a matrix of size $n \times \delta$. That is, we have a matrix for each variable and, consequently, k matrices of the same size, one for each indicator j , $j = \dots, k$. We interpret each column of the matrix as a realization of the j -th variable

over the units. That is, we have a random vector of normalized variable Z_j and we dispose of δ realization of this random vector.

To clarify this point, suppose we have an indicator that varies within the range $[10, 90]$. To normalize, we set $\delta = 11$ and $\epsilon = 1$, consequently $v = 0, 1, \dots, 10$. Thus, we can normalize the data by choosing numbers from 90 to 100 as the maximum and from 0 to 10 as the minimum. The values 0 and 100 are LB and UB , respectively. So, by choosing $v = 0$, we have $min = 10$ and $max = 90$, which corresponds to the classic min-max. When $v = 1$, $min = 9$ and $max = 91$. Therefore, we obtain a different normalized vector. We, thus, repeat the normalization process, generating 11 normalized vectors.

2.1 A simple proposal

As first proposal, we decide to define the *normalized value* for unit i as the mean value of the distribution of normalized indicators, \bar{Z}_j , and as measurement error the marginal error of the $(1 - \alpha)\%$ confidence interval for the mean value, ΔZ_j . That is,

$$\bar{Z}_1 \pm \Delta Z_1, \quad \bar{Z}_2 \pm \Delta Z_2, \quad \dots, \quad \bar{Z}_k \pm \Delta Z_k \quad j = 1, \dots, k. \quad (4)$$

We observe that ΔZ_j is the half-width of the confidence interval, that is the marginal error.

In the simplest case, we assume that, the composite indicator for each unit i is the arithmetic mean among the \bar{Z}_j ($j = 1, \dots, k$) indicators. We denote this quantity by $M = \frac{1}{k} \sum_{j=1}^k \bar{Z}_j$. However, as displayed in (4), each measure has an own confidence interval and we can take them into account. In this way we can affirm that two units are really different if their confidence intervals do not intersect.

To achieve this aim, we apply the classical physics law of propagation of uncertainty [6], working under the assumption of Independent Errors. More in detail, when the errors are independent, the uncertainty of the mean is: $\Delta M = \frac{1}{\sqrt{k}} \sqrt{\sum_{j=1}^k (\Delta Z_j)^2}$. Thus, the final result is:

$$M \pm \Delta M = \left(\frac{1}{k} \sum_{j=1}^k \bar{Z}_j \right) \pm \frac{1}{\sqrt{k}} \sqrt{\sum_{j=1}^k (\Delta Z_j)^2} \quad (5)$$

3. An illustrative example

3.1 Data

Data belong to the 35th edition (December, 2024) of the *Sole 24 Ore Quality of Life*.⁷

The Sole 24 Ore Quality of Life is an annual ranking of Italian provinces based on several indicators to capture the quality of life in those regions. It is published since 1990 by the Italian financial newspaper *Il Sole 24 Ore* and it is widely recognized as a comprehensive tool for assessing and comparing living conditions across the 107 Italian provinces.

⁷Data collected by the *Sole24ore* are available on the Sole Ore GitHub page by citizens, researchers, media and decision makers. <https://github.com/IlSole24Ore>

The ranking is based on six macro-categories and 90 indicators that capture different dimensions of well-being: i) Wealth and Consumption (it includes indicators such as per capita income, household savings, and property prices); ii) Business and Labor (key indicators include unemployment rates, the number of new businesses, and productivity levels); iii) Demography and Society (it includes population growth, aging, and migration rates); iv) Environment and Services (examples include air pollution levels, green spaces, public transportation, and waste management); v) Justice and Security (such as crime rates, road safety, and the efficiency of the judicial system), vi) Culture and Leisure (it includes the number of theaters, museums, sports facilities, and entertainment options).

Among the 90 indicators we focus on 3 variables related to environmental sustainability. More in details, we use 1) Urban ecosystem (Ecourb, Synthetic index on 18 parameters); 2) Electricity from renewable sources (Enrin, Incidence of wind, photovoltaic, geothermal and hydro, in % of gross production); 3) Protected Areas (Areeprot, %).

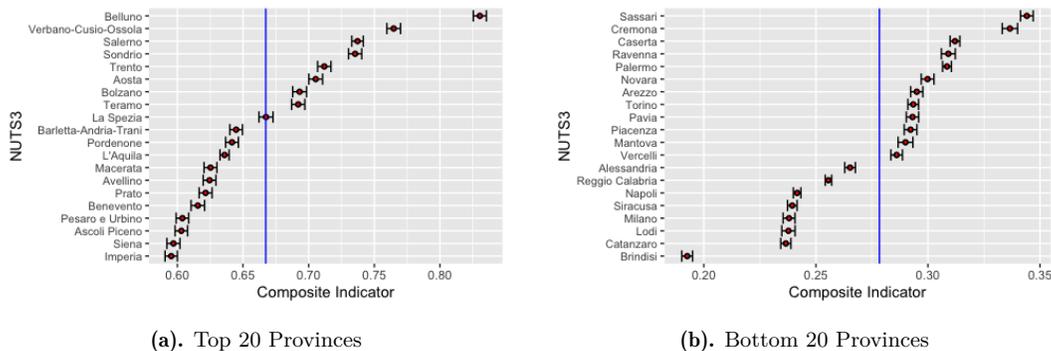
Table 1: Summary Statistics

	$\bar{Z}_1(\text{Ecourb})$	$\bar{Z}_2(\text{Enrin})$	$\bar{Z}_2(\text{Areeprot})$	M	LB(M)	UB(m)
Min.	0.01152	0.001802	0.002747	0.1926	0.1902	0.1950
Median	0.59423	0.495387	0.273077	0.4482	0.4438	0.4526
Mean	0.58248	0.501206	0.315993	0.4666	0.4629	0.4703
Max.	0.94300	0.951535	0.950706	0.8304	0.8254	0.8353

Notes. Our computation on Sole24ore data.

Figure 1 reports the 20 Provinces that occupy the Top position (left panel) and the Bottom positions (right panel). Blu line refers to the average value in the subset of Provinces. Belluno and Verbano-Cusio-Ossola occupy the first and second position, respectively. Their values are (statistically) different whereas Salerno and Sondrio can be ranked at the same position since their intervals intersect. In the bottom of the ranking we can find Brindisi which displays the worst performance whereas there is no differences among Catanzaro, Lodi, Milano, Sicarusa and Napoli.

Figure 1: Composite indicators with measurement errors



4. Conclusions

This work is a very preliminary attempt to account reliability on the normalization step by proposing to construct a distribution of normalized values. Future research will be directed towards involve error measurement according to different aggregation methods and to propose test on ranking differences.

Acknowledgment

This research has received funding from the research project *A survey-based Impact Evaluation of NRRP on Italian municipalities* funded by European Union - NextGenerationEU, Mission 4, Component 2, Investment 1.1, CUP I53D23007340001.

References

- [1] Matteo Mazziotta and Adriano Pareto. Everything you always wanted to know about normalization (but were afraid to ask). *Rivista Italiana di Economia Demografia e Statistica*, 75:41–52, 2021.
- [2] Joint Research Centre. *Handbook on constructing composite indicators: Methodology and user guide*. OECD publishing, 2008.
- [3] Francesca Mariani, Mariateresa Ciommi, Maria Cristina Recchioni, and Chiara Gigliarano. *A new point of view in the construction of composite indicators: a simulation illustration*. 2024.
- [4] Carlo Drago and Andrea Gatto. An interval-valued composite indicator for energy efficiency and green entrepreneurship. *Business Strategy and the Environment*, 31:2107–2126, 2022.
- [5] Francesca Mariani, Mariateresa Ciommi, and Maria Cristina Recchioni. A new class of composite indicators: The penalized power mean. *European Journal of Operational Research*, 317:1015–1035, 2024.
- [6] John R Taylor. *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. 1997.